DOCUMENT RESUME

ED 351 616 CG 024 616

AUTHOR	Onken, Lisa Simon; Blaine, Jack D.
TITLE	Psychotherapy and Counseling in the Treatment of Drug Abuse [Technical Review] (Rockville, Maryland, May 18-19, 1989). National Institute on Drug Abuse Research Monograph Series 104.
INSTITUTION	National Inst. on Drug Abuse (DHHS/PHS), Rockville, Md.
REPORT NO	ADM-90-1722
PUB DATE	90
NOTE	138p.
AVAILABLE FROM	National Clearinghouse for Alcohol and Drug Information, P.O. Box 2345, Rockville, MD 20852; Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402; or National Technical Information Service, U.S. Department of Commerce, Springfield, VA 22161 (free).
PUB TYPE	Books (010) -- Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Counseling Techniques; *Drug Abuse; *Drug Addiction; *Drug Rehabilitation; *Psychotherapy; Research Methodology; Substance Abuse

ABSTRACT

This monograph is based on the papers from a technical review. These papers are included: (1) Psychotherapy and Counseling Research in Drug Abuse Treatment: Questions, Problems, and Solutions (Lisa Onken, Jack Blaine); (2) Psychotherapy and Counseling for Methadone-Maintained Opiate Addicts: Results of Research Studies (George Woody, A. T. McLellan, Lester Luborsky, Charles O'Brien); (3) Crack and Cocaine Abusers in Outpatient Psychotherapy (Paula Kleinman, George Woody, Thomas Todd, Robert Millman, Sung-Yeon Kang, Jack Kemp, Douglas Lipton); (5) Therapist Effects in the Treatment of Drug Dependence: Implications for Conducting Comparative Treatment Studies (Paul Crits-Christoph, Katherine Beebe, Mary Beth Connolly); (6) Control Groups and Comparison Groups in Psychotherapy Outcome Research (T. D. Borkovec); (7) Attrition in Substance Abuse Comparative Treatment Research: The Illusion of Randomization (Kenneth Howard, W. M. Cox, Stephen Saunders); (8) Conceptualizing and Selecting Measures of Treatment Outcome: Implications for Drug Abuse Outcome Studies (Michael Lambert); (9) Can a Technology Model of Psychotherapy Research Be Applied to Cocaine Abuse Treatment? (Kathleen Carroll, Bruce Rounsaville); (10) Methodology: What Are the Design Issues Involved in the Defined Research Priorities? (Larry Beutler); and (11) Research Priorities for Psychotherapy and Counseling in the Treatment of Drug Abuse: The Psychotherapy Research Perspective (Perry London). (ABL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

National Institute on Drug Abuse

RESEARCH

MONOGRAPH SERIES

ED351616

Psychotherapy and Counseling in the Treatment of Drug Abuse

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CG024616

104

BEST COPY AVAILABLE

U.S. Department of Health and Human Services • Public Health Service • Alcohol, Drug Abuse, and Mental Health Administration

Psychotherapy and Counseling in the Treatment of Drug Abuse

Editors:

Lisa Simon Onken, Ph.D.
Division of Clinical Research
National Institute on Drug Abuse

Jack D. Blaine, M.D.
Division of Clinical Research
National Institute on Drug Abuse

NIDA Research Monograph 104
1990

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Alcohol, Drug Abuse, and Mental Health Administration

National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Psychotherapy and Counseling in the Treatment of Drug Abuse

ACKNOWLEDGMENT

This monograph is based on the papers from a technical review on "Psychotherapy and Counseling in the Treatment of Drug Abuse" held on May 18 and 19, 1989, in Rockville, MD. The review meeting was sponsored by the Office of Science and the Division of Clinical Research, National Institute on Drug Abuse.

COPYRIGHT STATUS

The National Institute on Drug Abuse has obtained permission from the copyright holders to reproduce certain previously published material as noted in the text. Further reproduction of this copyrighted material is permitted only as part of a reprinting of the entire publication or chapter. For any other use, the copyright holder's permission is required. All other material in this volume except quoted passages from copyrighted sources is in the public domain and may be used or reproduced without permission from the Institute or the authors. Citation of the source is appreciated.

Opinions expressed in this volume are those of the authors and do not necessarily reflect the opinions or official policy of the National Institute on Drug Abuse or any other part of the U.S. Department of Health and Human Services.

The U.S. Government does not endorse or favor any specific commercial product or company. Trade, proprietary, or company names appearing in this publication are used only because they are considered essential in the context of the studies reported herein.

DHHS publication number (ADM)90-1722
Printed 1990

NIDA Research Monographs are indexed in the *Index Medicus*. They are selectively included in the coverage of *American Statistics Index, BioSciences Information Service, Chemical Abstracts, Current Contents, Psychological Abstracts,* and *Psychopharmacology Abstracts.*

Contents

Page

Introduction

Psychotherapy and Counseling Research in Drug Abuse Treatment: Questions, Problems, and Solutions 1
Lisa Simon Onken and Jack D. Blaine

Research Findings

Psychotherapy and Counseling for Methadone-Maintained Opiate Addicts: Results of Research Studies 9
George E. Woody, A. Thomas McLellan, Lester Luborsky, and Charles P. O'Brien

Crack and Cocaine Abusers in Outpatient Psychotherapy 24
Paula H. Kleinman, George E. Woody, Thomas C. Todd, Robert B. Millman, Sung-Yeon Kang, Jack Kemp, and Douglas S. Lipton

Methodological Considerations

Therapist Effects in the Treatment of Drug Dependence: Implications for Conducting Comparative Treatment Studies 39
Paul Crits-Christoph, Katherine L. Beebe, and Mary Beth Connolly

Page

Control Groups and Comparison Groups in Psychotherapy Outcome Research 50
T.D. Borkovec

Attrition in Substance Abuse Comparative Treatment Research: The Illusion of Randomization 66
Kenneth I. Howard, W. Miles Cox, and Stephen M. Saunders

Conceptualizing and Selecting Measures of Treatment Outcome: Implications for Drug Abuse Outcome Studies 80
Michael J. Lambert

Can A Technology Model of Psychotherapy Research Be Applied to Cocaine Abuse Treatment? 91
Kathleen Carroll and Bruce Rounsaville

Methodology: What Are the Design Issues Involved in the Defined Research Priorities? 105
Larry E. Beutler

Research Priorities and Future Directions

Research Priorities for Psychotherapy and Counseling in the Treatment of Drug Abuse: The Psychotherapy Research Perspective 121
Perry London

List of NIDA Research Monographs 128

Introduction

Psychotherapy and Counseling Research in Drug Abuse Treatment: Questions, Problems, and Solutions

Lisa Simon Onken and Jack D. Blaine

Drug abuse treatment occurs in a multitude of forms. It may be provided in outpatient or inpatient settings, be publicly or privately funded, and may or may not involve the administration of medication. The differences among the philosophies of, and the services provided in, various drug abuse treatment programs may be enormous. What is remarkable is that some form of drug abuse counseling or psychotherapy is almost invariably a part of every type of comprehensive drug abuse treatment. Individual therapy or counseling is available in about 99 percent of the drug-free, methadone-maintenance, and multiple-modality drug abuse treatment units in this country (National Drug and Alcoholism Treatment Unit Survey 1982). It is also available in approximately 97 percent of the detoxification units.

Despite the fact that drug abuse counseling and psychotherapy are nearly universal in drug abuse treatment, surprisingly little is known about these forms of treatment. Much more research has focused on pharmacological treatments for drug abuse than on nonpharmacological, even though nonpharmacological interventions are almost always utilized and are sometimes the only form of treatment offered to the drug abuser.

In part, the paucity of research in this field is due to the inherent difficulties in scientifically investigating psychotherapy and counseling. It has not been uncommon for a psychotherapy/counseling research study to be denied funding because reviewers believed that fundamental, minimal standards for a scientific investigation had not been met. Even the most basic experimental standard, the double-blind method, is a virtual impossibility in comparative psychotherapy research.

There are two major dilemmas here: (1) Thousands of drug abusers are being treated every day with psychotherapy and counseling, and no one can say for certain what forms of psychotherapy and counseling are most effective, what types of therapists and counselors can best provide it, how long it should be provided, when psychotherapy is necessary, when counseling is sufficient, and so on, and (2) Doing research in this field is so "messy" that many a scientist would hesitate to attempt it, and those that do may end up discouraged by their difficulty in acquiring funds.

On May 18 and 19, 1989, a "Technical Review" was held in Rockville, MD, sponsored by the Treatment Research Branch, Division of Clinical Research, National Institute on Drug Abuse (NIDA), entitled "Psychotherapy and Counseling in the Treatment of Drug Abuse." Similar to many Technical Reviews that have been organized by NIDA, a major objective was to stimulate the psychotherapy and counseling research field by reviewing the research in that field and proposing future research directions. For the Technical Review on psychotherapy and counseling, two additional goals were considered to be of primary importance: (1) delineating the methodological problems in doing psychotherapy and counseling research and (2) describing the strategies researchers can use in dealing with those methodological problems.

The papers presented in that Technical Review are contained in this monograph and fall into three categories: (1) reviews of research findings; (2) methodological considerations; and (3) research priorities and conclusions.

Woody and colleagues provide an excellent summary of psychotherapy and counseling in the treatment of opiate abuse. In addition to reviewing the psychotherapy/counseling studies carried out by the University of Pennsylvania-Veterans Administration group, they present some interesting data from a current community-based project. The research done by this group represents some of the most solid and conclusive work done to date on psychotherapy and counseling in the treatment of methadone-maintained individuals. The results demonstrate the efficacy of psychotherapy (both dynamic and behavioral varieties), particularly for heroin addicts with moderate to high psychiatric severity.

Some interesting preliminary data on the retention of cocaine abusers in individual supportive-expressive therapy and structural-strategic family therapy are summarized in Kleinman and coworkers' contribution to this monograph. The most interesting, although preliminary, conclusion from their report is that the most potent predictor of retention is therapist assignment.

The second group of papers by Crits-Christoph and colleagues, Howard and colleagues, Borkovec, Lambert, Carroll and Rounsaville, and Beutler focuses

exclusively on methodological and design issues in carrying out psychotherapy and counseling research. Problems in doing this kind of research with drug abusers are emphasized wherever possible.

As London points out in the final chapter of this monograph, Crits-Christoph and coworkers focus on **therapist** variance; Borkovec addresses the issue of **therapy** variance; and Howard and coworkers look at the effects of attrition on **sample** variance. Numerous articles addressing therapist effects, in psychotherapy research studies with and without drug-abusing populations, are summarized in the Crits-Christoph article. A Monte Carlo study of therapists' effects is also presented, in which we are shown how wrong our conclusions can be about differences between therapies (or counseling strategies) if we ignore systematic differences between therapists (or counselors).

Borkovec outlines the advantages and disadvantages of research designs that attempt to determine: (1) differences in efficacy between therapies; (2) the effective components of therapies; and (3) mechanisms of change. A "perfect" research design in this field does not exist, and it is important to understand the benefits and limitations of any particular design. Borkovec does a superb job of pointing these out.

The related issues of randomization and attrition in psychotherapy/counseling research are the focus of the article by Howard and colleagues. They address the questions of how preinclusion and postinclusion attrition affects a sample and what, if anything, can be done to remedy these effects. An unfortunate conclusion that one must draw when reading this manuscript is that true "randomization" in a psychotherapy/counseling research study with drug abusers can never exist, given the reality of the attrition that occurs in these studies and given the fact that even if attrition rates are the same between groups, one cannot assume that the same types of subjects left the different groups. While some would argue that controlled random-assignment comparative psychotherapy/counseling research should be abandoned for this reason, most of the participants of the Technical Review agreed that this highlights the need for a variety of types of research, such as experimental or controlled, random-assignment studies and quasi-experimental or naturalistic studies. With the same questions being asked in a number of ways, convergent results among studies help to establish confidence in our answers.

At the conference, Carroll reviewed the literature on psychotherapy and counseling in the treatment of cocaine abuse and spoke about current research at Yale University. Studies in this area are few, and very little, if anything, can be stated conclusively with regard to this research. New, preliminary research at Yale, however, indicates that relapse prevention, a cognitive/behavioral strategy, may be helpful in the treatment of cocaine abuse (Carroll 1990). In this monograph, Carroll and Rounsaville discuss

the wisdom of applying a technology model to research in this field. Their article appears in this monograph with the other articles addressing methodological concerns.

Lambert's contribution to this monograph directs our attention to outcome measures. He points out the lack of consistency among psychotherapy research studies in types of outcome measures utilized, even among studies attempting to answer the same or similar questions. His article emphasizes the need for careful consideration when choosing these measures.

In the final chapter of the "Methodological Considerations" section, Beutler summarizes the methodological and design issues raised during the course of the conference. He also highlights those methodological issues most related to the research priorities defined by the group.

London provides an excellent synopsis and integration of the entire conference in the final contribution to this monograph. He also directs our attention to areas where further research is needed. He emphasizes the fact that although drug counselors are typically in positions of low pay and status, they are "on the treatment firing line" and that their work is extremely important. London believes that the study of drug abuse counseling and counselors, therefore, should be a primary focus of researchers in this field. London also suggests that due to the repeated assertion (in the absence of firm data) of the importance of 12-step programs as part of drug abuse treatment by clinicians and researchers in the drug abuse field, the secular and scientific study of such programs is essential.

It is our hope that both beginning and experienced drug abuse researchers in the psychotherapy/counseling field will benefit from the insights put forth in this monograph and not be too discouraged by the enormity of the problems intrinsic to this field. It is our belief that an understanding of these problems is the first step in dealing with them. While no one research study can overcome all of the problems described in this monograph, systematic research that addresses our questions from many different perspectives may provide us with the answers we need.

The Treatment Research Branch of NIDA has a strong interest in expanding and facilitating research on the psychotherapy and counseling of drug abusers. Following our Technical Review, we issued a program announcement entitled "Psychotherapy and Counseling in Drug Abuse Treatment" to solicit applications for the funding of original research studies in this area. We have also initiated the development of a national, multisite collaborative study to maximize and compare the efficacy of psychodynamic psychotherapy, cognitive/behavioral therapy, and drug counseling in the treatment of cocaine abusers.

It is our belief that the importance of the "talking" therapies in the treatment of drug abuse cannot be underestimated. Even where effective pharmacotherapies are already available for drug abuse treatment, few would argue that they should be administered without concurrent counseling or psychotherapy except under emergency circumstances when it is not possible to provide these services. Many would argue that pharmacotherapy should **never** be given in isolation. In any event, effective pharmacotherapies do not exist for every form of substance abuse. The mainstay of drug abuse treatment is psychotherapy and/or drug counseling. It is crucial that we understand how and why they work.

REFERENCES

Carroll, K. Relapse prevention and interpersonal psychotherapy for ambulatory cocaine abuse. Presented at the 143rd Annual Meeting of the American Psychiatric Association, New York, May 12-17, 1990.

National Drug and Alcoholism Treatment Unit Survey. DHHS Pub. No. (ADM)89-1626. Rockville, MD: U.S. Department of Health and Human Services, 1982.

AUTHORS

Lisa Simon Onken, Ph.D.
Research Psychologist
Jack D. Blaine, M.D.
Chief

Treatment Research Branch
Division of Clinical Research
National Institute on Drug Abuse
Parklawn Building, Room 10A-30
5600 Fishers Lane
Rockville, MD 20857

Research Findings

Psychotherapy and Counseling for Methadone-Maintained Opiate Addicts: Results of Research Studies

George E. Woody, A. Thomas McLellan, Lester Luborsky, and Charles P. O'Brien

INTRODUCTION

The data presented in this chapter deal with psychotherapy for methadone-maintained opiate addicts. Data from a psychotherapy study completed several years ago in the methadone treatment program at the Philadelphia Veterans Administration Medical Center are reviewed, and preliminary findings from a related study that is currently under way in three community-based methadone programs are presented. The part of this chapter dealing with the first study will be familiar to many; thus, its rationale, background, and major findings are briefly summarized.

Interestingly, the impetus for research by the National Institute on Drug Abuse (NIDA) on psychotherapy with methadone patients was driven more by criticism of the outcomes being achieved by addicts in maintenance treatment than it was by enthusiasm for psychotherapy. The early results with methadone were excellent (Dole and Nyswander 1965; Dole et al. 1969). Patients experienced a marked reduction in illicit drug use and criminal behavior, and many became regularly employed for the first times in their lives. Patients who entered maintenance treatment in these early studies were carefully selected as addicted only to heroin and relatively free of complicating medical or psychiatric problems. The staff was highly motivated and trained, and there was much enthusiasm among patients, as all were volunteers for this new and exciting treatment. These conditions were such that the best possible outcomes with methadone were probably attainable in these early studies.

As a result of the generally excellent outcomes with patients treated under these special conditions, methadone was rapidly disseminated throughout the country, and currently more than 80,000 patients are in maintenance treatment. The programs were often developed with great haste in response to a very serious heroin problem, and most staff had little training or experience with opiate addicts or methadone treatment. The bulk of the "on-line" treatment was done by paraprofessionals, many of whom had only bachelor's degrees or less. Some were ex-addicts with considerable personal experience but less than a high school education. Salaries were often low. An example of the staff recruitment efforts is seen below in a recent advertisement from the "Help Wanted" section of the *Philadelphia Inquirer*. It is therefore not surprising that counselors hired from this ad would be poorly equipped to deal with the complex psychosocial problems presented by the heroin addict.

Philadelphia Inquirer
Sunday, April 10, 1988

COUNSELOR for D/A Prog.
Full Time. No Experience
Necessary. $12,000. Send
resume to: XXX XXXXX XX
Phila., PA.

In addition to the potential problems in treatment delivery caused by inexperienced and untrained staff, many of the patients started on methadone were abusing other drugs in addition to heroin, while others had a range of behavioral and psychiatric problems that were not always easily managed.

Perhaps as a consequence of these differences in staff training and patient selection, the results obtained by many programs were less positive than those found in the experiments of Dole and Nyswander 1965. Some critics even said that methadone was ineffective, a claim found to be untrue in later studies. One of the criticisms often voiced was that the addicts being treated with methadone were some of the most disturbed individuals in the mental health care system, but that the people providing their day-to-day treatment were some of the least trained. Specifically, clinicians noted that many addicts had psychiatric problems that could benefit from additional psychiatric treatment. This treatment was usually not available because the programs had few employees with the training and skills necessary for the delivery of these services.

As a result of these criticisms, NIDA organized a series of meetings that were designed to develop testable hypotheses about the interface between psychopathology and addiction (Blaine and Julius 1977). Two lines of research were proposed. One was the evaluation of methadone patients

psychiatrically and the determination of the types and frequencies of psychiatric disorders that they experienced, both currently and over their lifetimes. The second research area was the determination of whether any evidence that professional psychotherapy can be helpful to methadone patients exists.

PART I—THE FIRST SERIES OF STUDIES

Three studies were funded. One was a psychiatric diagnostic study done in Boston by Khantzian and Treece (1985). The second was a comprehensive project that contained both a diagnostic study and a psychotherapy study, carried out by Rounsaville and coworkers (1983) in New Haven at Yale University. The third was a psychotherapy study done at the University of Pennsylvania/ Philadelphia Veterans Administration (VA) Medical Center (Woody et al. 1983). This project also had a diagnostic component, as all patients entered into the study had a psychiatric evaluation. Several measures were used at all treatment sites for purposes of increasing the chances for comparability of data. The most comprehensive of these was the SADS-L interview, which was used to make DSM-III and RDC diagnoses.

The details and results of each study have been described in a series of published reports. The results of the diagnostic studies were almost identical. All found that 80 to 85 percent of the methadone patients had a range of psychiatric disorders in addition to opiate dependence, either at the time of the interview or in the past. The most common, occurring in 50 to 60 percent of the samples, were depressive disorders, usually major depression. Antisocial personality disorder was found in approximately 20 to 50 percent of each sample, depending on whether the RDC or DSM-III criteria were used. Alcohol dependence, either current or past, was found in 20 to 25 percent; anxiety disorders in 10 to 20 percent; and an assortment of other problems, often reflecting disorders of mood, such as labile personality or bipolar II disorder, were found in 2 to 10 percent (Rounsaville et al. 1982). The New Haven study also evaluated a group of addicts who were not in treatment and found the same types of problems as seen in the treated sample; however, the out-of-treatment subjects overall had fewer disorders than did the in-treatment group (Rounsaville and Kleber 1985). One interpretation of this finding was that coexisting psychiatric problems may have contributed to the decision to enter treatment.

The psychotherapy studies achieved differing results. The Yale group compared drug counseling plus interpersonal psychotherapy (IPT) with counseling alone and found that all patients improved but that there were no differences between groups (Rounsaville et al. 1983). The VA/Pennsylvania study compared counseling plus supportive expressive therapy (SE) or cognitive behavioral therapy (CB) with counseling alone (DC) and found that all patients improved but that those who received the extra

psychotherapy achieved better outcomes in more areas than did those who received counseling alone (Woody et al. 1983).

Possible reasons for the differences in outcome achieved in these two studies have been discussed elsewhere and include: differences in acceptability and motivation by patients between the sites (the New Haven study had a relatively high dropout rate and a lower overall number of subjects than did the VA/Pennsylvania study); differences in administrative structure for running the study; differences in efficacy between therapists at the two sites; and better overall efficacy of the drug-counseling program at the New Haven site (urine test results were "cleaner" than at the VA/Pennsylvania site).

The greater number of subjects who entered and completed the protocol at the VA/Pennsylvania site (110 vs. 73 at New Haven) provided an opportunity to examine interactions between patient types, therapy, and therapist. The data obtained provide interesting and potentially meaningful guides for determining which types of patients may benefit from psychotherapy and also which therapist qualities are associated with positive outcomes. These interactions will be summarized below as they suggest how psychotherapy might be applied to drug-treatment programs.

Psychiatric Severity

The first interaction seen was that between psychiatric severity and outcome. Previous studies done at the VA/Pennsylvania program had shown that a global rating of psychiatric severity is the best predictor of outcome for both opiate addicts and alcoholics being treated in a range of outpatient and inpatient programs. This work showed that patients with few additional psychiatric symptoms (termed low-severity patients) generally did well in all programs. Patients with high symptom levels generally did poorly in any program. Midseverity patients had intermediate outcomes and appeared to be the group that was most sensitive to patient/program matches (McLellan et al. 1983).

The psychotherapy study data were examined according to psychiatric severity, especially for interactions among severity, outcome, and treatment condition. It was found that there were few differences in outcome between groups of low-severity patients among the three treatment conditions (SE, CB, and DC). High-severity patients who received psychotherapy showed a number of gains, but little progress was made if they received drug counseling alone. Midseverity patients showed more gains with psychotherapy than with counseling alone, but patients in each treatment condition improved in several areas. The conclusion was that the addition of psychotherapy altered the traditional relationship between high-psychiatric severity and poor outcome. Psychotherapy gave this group of more

disturbed patients a better chance to benefit from methadone treatment (Woody et al. 1984).

This finding pointed toward a possible cost-effective use of psychotherapy. High-severity patients are a very difficult group to treat. They usually demand more staff time than do their less disturbed counterparts and make little progress. The study results imply that these patients can be identified early in treatment, provided additional psychotherapy, and thus given a better chance to improve. This plan could also reduce the strain and time demands that these patients place on drug counselors (Woody et al. 1986).

Antisocial Personality Disorder

The second interaction examined was that between antisocial personality disorder (ASPD) and outcome. Many opiate addicts have ASPD, and people with this diagnosis typically do not respond well to treatment. A literature review, however, indicated that there are probably many subtypes of ASPD and that some patients with this diagnosis may be "therapy responsive." With this in mind, we examined those with ASPD and found that about half had other Axis I diagnoses, most commonly depression. We then examined four groups of patients who received psychotherapy: (1) those with a diagnosis of opiate dependence only; (2) those with opiate dependence and depression; (3) those with opiate dependence, depression, and ASPD; and (4) those with opiate dependence and ASPD only.

We found that patients in groups 1 and 2 showed gains in many areas, especially those in group 2. Patients in group 3 showed considerable progress also, but not quite as much as those in the first two groups. In contrast, patients with only opiate dependence and ASPD (group 4) showed gains in only a few measures of drug use but no significant changes in any of the other areas measured (Woody et al. 1985). Thus, this analysis confirmed the impression that ASPD is a negative predictor of outcome, but it also indicated that patients with depression accompanying their ASPD can respond to therapy. One possible explanation is that those with depression have more capacity to relate to people and events and to experience feelings such as guilt or loss; another is that depression is a psychiatric problem that is responsive to psychotherapy and that patients with ASPD and depression respond simply because they happen to have an associated condition that is amenable to treatment.

Therapist Assignment

The third analysis of patient/therapy interactions examined outcome according to therapist assignment. Psychotherapy studies have traditionally examined outcome according to treatment assignment. There have been attempts to examine the qualities that are associated with successful outcome, but most studies have paid little attention to examining the

interaction between therapist assignment and outcome within a specific treatment modality. This study had employed five SE and four CB therapists, and we were able to determine whether therapist assignment and outcome were related. We chose three therapists and counselors from each modality who had treated at least seven patients, and we compared the overall outcomes of patients in their caseloads.

As seen in table 1, there were significant differences in outcome, as judged by the effect size. One SE therapist had a very large effect, while another had very little effect and in some cases may have made patients worse. Similar but less dramatic variability in outcomes was seen for CB therapists and drug counselors. Similar variability in outcome according to therapist assignment was found by Luborsky and coworkers (1986) in an analysis of results from other psychotherapy studies and by McLellan and coworkers (1988) in a study of outcome according to counselor assignment.

The next step was to identify the factors that contributed to these differences in outcome. We had tape-recorded most of the therapy sessions, and sections of the tapes were rated by two independent raters according to how closely the therapists and counselors conformed to the specifications of their respective treatment manuals. We also had asked the patients to fill out a form after completing their third session, indicating the degree to which they felt that the therapist or counselor was helping them. These ratings of "helping relationship" and compliance to the specifications of therapy were examined to determine whether either one predicted outcome. The strongest predictor of outcome was the patient's rating of the helping relationship. The degree to which the therapist conformed to the specifications of the therapy technique also predicted outcome, but to a smaller degree than the relationship variable. One possible interpretation of this finding is that the relationship is a necessary factor for the technique to work and without a good relationship, any technique that is applied will not be used productively by the patient. Another possibility is that superior therapists do better in many ways, including building positive relationships and complying with their specified techniques.

Throughout all analyses, both SE and CB patients generally were associated with similar amounts of improvement, and thus we found no advantage for one therapy over the other with this population. The differences in outcome between SE, CB, and DC patients that were seen at 7 months were also seen at the 12-month followup, 6 months after therapy ended (Woody et al. 1987).

In brief, our experience with this study showed that the additional therapy could provide meaningful benefits to opiate-addicted veterans being treated in the methadone program, particularly those with significant psychiatric symptoms in addition to the addiction. Interestingly, the drug counselors

TABLE 1. *Percent change from start of treatment to 7-month followup*

Therapists	Number	Outcome Measures[a]			
		ASI			
		Drug Use	Employment Status	Legal Status	Psychiatric Status
SE					
A	10	34	32	20	82
B	8	33	34	17	41
C	8	-14	12	7	-1
CB					
D	11	61	19	17	36
E	10	70	22	13	17
F	9	48	10	11	14
DC					
G	9	51	8	13	7
H	6	46	-4	6	10
I	7	66	17	7	12

Therapists	Number	Outcome Measures[a]			
		Beck Depression Scales	SCL-90	Maudsley Scale	Average Effect Size[b]
SE					
A	10	58	44	64	0.74
B	8	37	46	59	0.59
C	8	8	-2	13	0.19
CB					
D	11	36	39	44	0.53
E	10	24	30	30	0.44
F	9	14	21	33	0.46
DC					
G	9	4	9	-1	0.20
H	6	-3	11	3	0.13
I	7	14	15	17	0.27

NOTE: ASI, Addiction Severity Index; SCL-90, Hopkins Symptom Checklist-90; SE, supportive-expressive; CB, cognitive-behavioral; and DC, drug counseling.

[a]All criteria were measured during the 30 days before treatment start and before 7-month followup. Factor scores represent composites of several items indicative of patient status in that area. Percent change was calculated against the treatment-start baseline.

[b]Within-therapist effect size was averaged across all seven criteria. Effect-size calculation for each criterion was pretreatment mean minus posttreatment mean, divided by pretreatment mean, divided by pretreatment SD. Small change=0.2; moderate change=0.5; large change=0.8.[36]

SOURCE: Luborsky et al. 1985, copyright 1985, *Archives of General Psychiatry*.

20

began their participation in the study with feelings of reluctance, competition, and uneasiness, but after working with the protocol for several years, most became very active in identifying possible study candidates. They were especially active in referring those with additional psychiatric symptoms in hopes that random assignment would place the patients they recruited in one of the therapy groups. We interpreted this change in attitude to the help provided by the therapists with the more disturbed patients and viewed this as an experiential confirmation of the data analyses. More detail about the techniques used and the overall results is summarized in two recent publications (Woody et al. 1986; Woody 1989).

It is important to note that this work was really a combined pharmacotherapy/counseling/psychotherapy study. All of these ingredients were necessary to perform the work. The patients would not have been available for therapy without the methadone; the concrete services and drug-focused therapy provided by the counselors helped manage the addiction and many of the associated social problems; and the psychotherapists provided additional help for those with the more complicated psychiatric problems.

PART II—CURRENT WORK: EXPERIENCES WITH COMMUNITY-BASED PROGRAMS

The next step was to see what would happen if we attempted a similar project in community-based programs. Our treatment program is very different from most methadone programs, mainly in the amount of staff and services that are available but also in the number of research projects going on at any one time. The patients are all veterans and almost entirely male and thus are not identical to those being treated in community-based programs where 25 to 30 percent are female, about 10 percent are veterans, and the overall educational levels are 1 to 2 years lower.

The design chosen for this work was somewhat different from that of the original project. We selected only one therapy, SE, rather than two. This was done because we found nonsignificant differences in comparisons between the SE and CB. Also, the fact that SE therapists are easier to find (more people have been trained in SE than in CB, even in Philadelphia) was a consideration. We used the same outcome measures at the same intervals as in the first study, but we randomly assigned patients to counseling plus SE (the SE group) or to a two-counselor group (the DC/DC group). This was a more conservative design and was done to control better for time and also for the possibility that the differences seen were attributable to the fact that patients receiving psychotherapy had a better chance to improve because they had a chance to form "helping" relationships with two individuals rather than one. A final modification was choosing only those patients who had moderate-to-high levels of psychiatric symptoms. This was done because the earlier work showed that there was

little advantage in providing those having low symptom levels with additional therapy.

We began the project in two community programs. At the end of the first year, we terminated our relationship with one of these programs because it was clear that it was not capable of supplying the number of patients necessary to complete the project. This program was replaced with two others, making a total of three. We have continued to work with each of these three programs and hope to finish the project with two therapists having treated at least 10 to 12 patients in each program, with 15 to 20 DC/DC controls/program. This should permit us to analyze the data by therapist and by program, in addition to the overall results.

PRELIMINARY FINDINGS

Recruitment and Engagement

It is more difficult to carry out the project in these programs than in our own. Perhaps this is because we have less control over the staff and less ability to prioritize the recruitment of subjects among the competing time demands facing the clinic staffs. In the VA study, about 80 percent of those who met inclusion criteria and were asked to join the study agreed to do so, and about 80 percent of these kept their three initial appointments. In the current study, approximately 60 percent of those who meet criteria agree to participate, and about 60 percent of those who complete study intake keep the three initial appointments. These data are summarized in table 2 below.

TABLE 2. *Compliance according to clinic*

Clinic	Started	Engaged	Percentage Engaged
A	63	43	68
B	14	7	50
C	35	16	46
D	35	26	74

As seen, there are differences in rates of engagement between clinics, and we think that these are related to the quality of staff and to clinic organization. For example, clinic D has the highest educational level among its staff (most have master's degrees), the best organization, and a relatively low staff turnover. Clinic C has had administrative problems and a 200- to 300-percent staff turnover since the project began. Clinic B is the program where work was discontinued due to its inability to supply an adequate number of subjects.

TABLE 3. *Psychotherapy study—SE vs. DC, 7-month followup*[a]

Variable	SE Baseline (n=39)	t	SE Followup (n=39)	ANCOVA *P*
Medical Factor	41		28	
Days Medical Problems	12	+	7	
Employment Factor	72		68	
Days Worked in Past 30	3	**	8	
Employment Income	$207.00	*	$470.00	
Welfare Income	$161.00		$181.00	
Drug Factor	34	**	29	
Days Opiate Use	8		7	
Days Stimulant Use	4		3	
Days Depressant Use	11		9	
Legal Factor	13		15	
Days Illegal Activity	3		4	
Illegal Income	$347.00		$124.00	
Psychiatric Factor	41	*	33	
Days Psychological Problems	15	**	9	
Beck	21	**	15	
SCL	107	**	83	
Maudsley	32		31	

Variable	DC Baseline (n=17)	t	DC Followup (n=17)	ANCOVA *P*
Medical Factor	49		39	
Days Medical Problems	10		12	+
Employment Factor	75		72	
Days Worked in Past 30	7		8	
Employment Income	$308.00		$334.00	
Welfare Income	$161.00		$214.00	
Drug Factor	40	+	33	
Days Opiate Use	12		6	
Days Stimulant Use	9		5	
Days Depressant Use	9		6	
Legal Factor	16	***	8	
Days Illegal Activity	5		2	
Illegal Income	$297.00		$328.00	
Psychiatric Factor	42	+	30	
Days Psychological Problems	18		13	
Beck	25	*	16	
SCL	111	+	77	
Maudsley	31	+	27	

+P<.08 (trend); *p<.05; **p<.01; ***p<.001

[a]April 1989

TABLE 4. *Psychotherapy study—DC vs. SE, 12-month followup*

Variable	Baseline (n=36)	t	12-Month Followup (n=36)	ANCOVA *P*
		SE		
Employment Factor	72	*	65	
Days Worked in Past 30	3	**	8	
Employment Income	$169.00		$352.00	
Welfare Income	$169.00		$223.00	
Drug Factor	33	***	24	
Days Opiate Use	8		6	
Days Stimulant Use	4		2	
Days Depressant Use	10		9	
Legal Factor	8		9	
Days Illegal Activity	2		2	
Illegal Income	$85.00		$67.00	
Psychiatric Factor	42	***	28	
Days Psychological Problems	15	***	8	
Beck	21	***	11	
SCL	108	*	87	
		DC		
Variable	Baseline (n=15)	t	12-Month Followup (n=15)	ANCOVA *P*
Employment Factor	71		79	+
Days Worked in Past 30	8		5	
Employment Income	$349.00		$245.00	
Welfare Income	$136.00		$183.00	
Drug Factor	41	*	30	
Days Opiate Use	14	*	4	
Days Stimulant Use	9		7	+
Days Depressant Use	10		10	
Legal Factor	18		11	
Days Illegal Activity	6	*	3	
Illegal Income	$817.00		$135.00	
Psychiatric Factor	42		35	
Days Psychological Problems	18		13	
Beck	26	**	15	
SCL	108		93	

+P<.08 (trend); *p<.05; **p<.01; ***p<.001

Outcome

Tables 3 and 4 summarize 7- and 12-month followup data. The numbers are relatively small, and the data are overly representative of the work of one therapist in one clinic. Despite these caveats and the fact that the results might change with more therapists, it appears that both groups are making progress but that the additional psychotherapy is providing more gains than in the DC/DC group, especially at the 12-month followup point.

The differences between the groups at 7 months are not as great as those seen in the earlier study, but the design is more conservative because of the addition of a second counselor to the DC group.

Therapist Differences

We do not yet have sufficient patients to do meaningful analyses of outcome by therapist. We have data on engagement rates, and it appears that there are significant differences between therapists at this level. For example, one therapist engaged approximately one-third of the assigned patients, and another engaged two-thirds. We hope to explore these differences more fully in future analyses, as was done in the first project.

Clinic Differences

These differences have been mentioned above in discussions of factors that may relate to the differences in engagement rates between clinics. After work on this project had been ongoing for about 1 year, it became apparent that there were major differences in structure, organization, and patient outcome between clinics. This finding was totally unexpected and particularly notable because there do **not** appear to be major differences between the patients who are treated at each of these community-based programs. It simply appears that some clinics do better than others. Figure 1 illustrates this point.

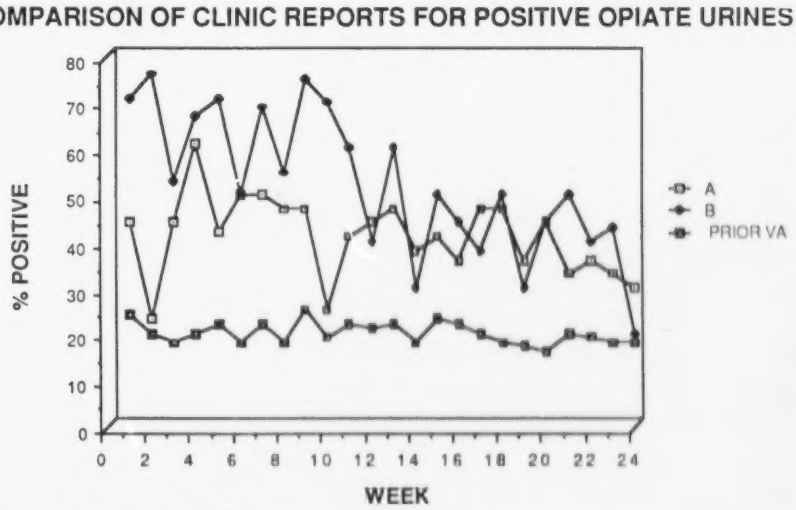

FIGURE 1. *Comparison of clinic reports for positive opiate urines*

As seen in figure 1, though the patients are reasonably comparable, the urine-test results for positive opiates are quite different. This finding is similar to that reported recently by Ball and coworkers (1989) in a project studying outcome and clinic structure in seven different methadone programs. We intend to explore these differences in greater depth toward the end of the project and are particularly interested in how they may relate to the acceptance of psychotherapy and to patient outcome.

SUMMARY

Data have been presented about the potential role of psychotherapy for psychiatrically impaired methadone-maintained opiate addicts. Complete data from one study, and preliminary data from a second, indicate that professional psychotherapy can be helpful as a supplement to ongoing drug-counseling services for patients having clinically significant psychiatric symptoms. If psychotherapy is to be used, care must be taken to integrate it into the ongoing clinical services of the methadone-treatment program.

Not all therapists are equally adept at engaging and working with addicts. In hiring therapists, attempts should be made to identify those who are not only technically competent but also interested and comfortable with this population. It should also be emphasized that there is considerable variability among methadone programs in such vital areas as leadership, staffing patterns, organization, dosing procedures, location, physical plant, and availability of ancillary services. These administrative differences may play a significant role in the feasibility and success of attempts to use psychotherapy in drug-treatment programs.

Finally, it should be noted that there is no evidence that psychotherapy cures addiction or that it can be used successfully without integrating it into other important services, such as drug counseling, methadone treatment, and the overall program structure. There is reason to believe, however, that it can provide additional and clinically meaningful benefits to that subgroup of methadone patients who are psychiatrically impaired.

REFERENCES

Ball, J.C.; Lange, W.R.; Myers, C.P.; and Friedman, S.R. Reducing the risk of AIDS through methadone maintenance treatment. *J Health Soc Behav* 29:214-226, 1989.

Blaine, J.D., and Julius, D.A., eds. *Psychodynamics of Drug Dependence.* National Institute on Drug Abuse Research Monograph 12. U.S. Department of Health, Education, and Welfare, 1977.

Dole, V.P., and Nyswander, M. A medical treatment for diacetylmorphine (heroin) addiction. *JAMA* 193(8):646-650, 1965.

Dole, V.P.; Robinson, J.W.; Orraca, J.; Towns, E.; Seargy, P.; and Caine, E. Methadone treatment of randomly selected criminal addicts. *N Engl J Med* 280:1372-1375, 1969.

Khantzian, E.J., and Treece, C. DSM-III psychiatric diagnoses of narcotic addicts: Recent findings. *Arch Gen Psychiatry* 42:1067-1071, 1985.

Luborsky, L.; Crits-Christoph, P.; McLellan, A.T.; and Woody, G.E. Do psychotherapists vary in their effectiveness? The answer within four outcome studies. *Am J Orthopsychiatry* 56(4), 1986.

Luborsky, L.; McLellan, A.T.; Woody, G.E.; O'Brien, C.P.; and Auerbach, A. Therapist success and its determinants. *Arch Gen Psychiatry* 42:602-611, 1985.

McLellan, A.T.; Luborsky, L.; Woody, G.E.; Druley, K.A.; and O'Brien, C.P. Predicting response to alcohol and drug abuse treatments: Role of psychiatric severity. *Arch Gen Psychiatry* 40:620-625, 1983.

McLellan, A.T.; Woody, G.E.; Luborsky, L.; and O'Brien, C.P. Is the counselor an "active ingredient" in substance abuse treatment? *J Nerv Ment Dis* 176(7):423-430, 1988.

Rounsaville, B.J.; Glazer, W.; Wilbur, C.H.; Weissman, M.M.; and Kleber, H.D. Short-term interpersonal psychotherapy in methadone-maintained opiate addicts. *Arch Gen Psychiatry* 40:630-636, 1983.

Rounsaville, B.J., and Kleber, H.D. Untreated opiate addicts: How do they differ from those seeking treatment? *Arch Gen Psychiatry* 42:1072-1077, 1985.

Rounsaville, B.J.; Weissman, M.M.; Kleber, H.; and Wilber, C. The heterogeneity of psychiatric diagnosis in treated opiate addicts. *Arch Gen Psychiatry* 39:161-166, 1982.

Woody, G.E. Psychotherapy for opiate dependence. In: Kleber, H., ed. *American Psychiatric Association Handbook on Psychiatric Treatments.* Washington, DC: American Psychiatric Association Press, 1989.

Woody, G.E.; Luborsky, L.; McLellan, A.T.; O'Brien, C.P.; Beck, A.T.; Blaine, J.; Herman, I.; and Hole, A. Psychotherapy for opiate addicts: Does it help? *Arch Gen Psychiatry* 40:639-645, 1983.

Woody, G.E.; McLellan, A.T.; Luborsky, L.; and O'Brien, C.P. Sociopathy and psychotherapy outcome. *Arch Gen Psychiatry* 42:1081-1086, 1985.

Woody, G.E.; McLellan, A.T.; Luborsky, L.; and O'Brien, C.P. 12-month follow-up of psychotherapy for opiate addiction. *Am J Psychiatry*, 144(5):590-596, 1987.

Woody, G.E.; McLellan, A.T.; Luborsky, L.; O'Brien, C.P.; Blaine, J.; Fox, S.; Herman, I.; and Beck, A.T. Psychiatric severity as a predictor of benefits from psychotherapy: The Penn-VA study. *Am J Psychiatry* 141:10, 1172-1177, 1984.

Woody, G.E.; O'Brien, C.P.; McLellan, A.T.; and Luborsky, L. Psychotherapy as an adjunct to methadone treatment. In: Meyer, R., ed. *Psychiatric Aspects of Opiate Dependence.* New York: Guilford Press, 1986. pp. 169-195.

AUTHORS

George E. Woody, M.D.
A. Thomas McLellan, Ph.D.
Lester Luborsky, Ph.D.
Charles P. O'Brien, M.D., Ph.D.

Substance Abuse Treatment Unit
Veterans' Administration Medical Center
Building 17
University and Woodland Avenues
Philadelphia, PA 11904

Crack and Cocaine Abusers in Outpatient Psychotherapy

Paula H. Kleinman, George E. Woody, Thomas C. Todd, Robert B. Millman, Sung-Yeon Kang, Jack Kemp, and Douglas S. Lipton

INTRODUCTION

Background

Cocaine and crack are the source of more professional and public concern than is any other substance on the drug scene today. Cocaine is not only the sole illicit substance whose use has been on the increase throughout the 1980s (Schuster 1988), it is also a drug for which treatment is still in its infancy (Kleber and Gawin 1984; Rosecan et al. 1987; Washton and Gold 1987). Many treatments are being tried, and it is important to learn even preliminary lessons about treatments for cocaine addiction. Psychosocial approaches are part of almost all treatments for cocaine abuse, and thus study of these therapies are indicated and appropriate.

Family Therapy and Psychotherapy as Treatments for Drug Addiction

The theoretical premises underlying structural-strategic family therapy (Haley 1980; Stanton et al. 1982; Todd et al. 1985) and supportive-expressive (SE) individual therapy (Luborsky et al. 1977) suggest that both might be effective treatments for cocaine abuse.

Family therapy maintains that the family must be approached as a system (Bateson et al. 1956; Haley 1973; Hoffman 1981) and that family behavior patterns contribute to maintaining the symptoms of its "problem member." In this model, the symptoms of the identified patient function to maintain a balance or homeostasis within the family (Jackson 1957; Stanton et al. 1982), and the family dynamics underlying the need for one member to become symptomatic must be identified and worked with before the "patient" can improve.

SE therapy aims to reduce drug use by altering problematic relationship patterns or psychological symptoms that contribute to drug involvement. Its focus is on increasing the patient's understanding of the meanings of the drug abuse, especially the stresses that precipitate and maintain the drug dependence, and of core conflicts that occur in relationships with others. The supportive component is designed to stabilize the patient, strengthen his/her defenses, prevent regression, and increase the patient's ability to benefit from the expressive aspect of the treatment. The SE treatment may thus decrease drug use by reducing underlying symptoms that contribute to the drug use.

Previous work has shown that both family therapy and SE therapy combined with methadone maintenance are more effective than methadone maintenance and paraprofessional counselors alone, especially for psychiatrically symptomatic patients (Stanton and Todd 1982; Woody et al. 1983).

Overview

In view of the data from methadone treatment, the authors undertook a study whose objective is to test the effectiveness of three different forms of therapy for serious cocaine abuse: SE therapy, family therapy, and paraprofessionally led group therapy. Both SE therapy and family therapy were manually guided. The group was eclectically run, with a focus on drug use, current problems (especially as they were revealed in ongoing group dynamics), and relapse prevention. It was not manually guided and was designed as a "treatment-as-usual" or an active control-group condition. Clients eligible for inclusion in the study were required to have either a spouse or parents willing to participate in treatment. They were selected for initial interview from a larger pool of persons who called a cocaine treatment hotline. At the conclusion of the initial screening interview, if willing to participate in the study, the client was assigned on a random basis to one of the three treatment conditions. In all, 148 individuals or families meeting these criteria started in the project between June 1987 and November 1988. Of these clients, 80 percent (117) were members of a spousal family, and the remaining 20 percent (31) were members of a parental family.

This chapter reports on the sociodemographic characteristics of the clients, describes client psychopathology in terms of DSM-III diagnoses, and presents data on predictors of retention. Finally, preliminary data on cocaine use during treatment and preliminary data on cocaine use at the time of 6-month followup are presented.

MEASURES

All subjects received an intake battery; data used in this chapter were drawn from a subset of the whole, including the Structured Clinical Interview for

30

BEST COPY AVAILABLE

DSM-III-R (SCID), the Symptom Checklist-90 (SCL-90), the Addiction Severity Index (ASI), and a drug history instrument designed by the authors. In addition to the preceding, the entire battery included the Beck Depression Inventory (Beck and Steer 1981), FACES III (Olson et al. 1985), a videotaped family task coded according to dimensions described by Stanton and coworkers (1982), and the Shipley Institute of Living Scale (Nunnally 1967). It took two and one-half to three and one-half hours to complete and was administered over the course of two sessions.

SCL-90

The SCL-90 is a self-administered measure of symptomatic distress, on which respondents rate themselves on a scale ranging from 0 to 4 for each of the 90 component items. It is intended to reflect underlying disturbances in the following nine areas: (1) somatization, (2) obsessive-compulsivity, (3) interpersonal sensitivity, (4) depression, (5) anxiety, (6) hostility, (7) phobic anxiety, (8) paranoid ideation, and (9) psychoticism (Derogatis 1983).

SCID

The SCID, Outpatient Version (SCID-OP), and the SCID, Personality Disorders (SCID-II), are relatively new, interviewer-administered schedules designed to diagnose DSM-III-R Axis I and Axis II disorders. They were developed by Spitzer and coworkers (1987) and designed to improve upon the Schedule for Affective Disorders (SADS). The SCID-OP and SCID II are relatively new measures, and only preliminary reliability studies have been reported to date. Reliability of the diagnosis of cocaine dependence was reported to be k=1.0, and for major depression, preliminary results were reported to be k=.60 (Gibbon 1987).

Addiction Severity Index

The Addiction Severity Index (ASI) is a structured clinical research interview that measures the severity of existing problems in the following areas: medical, legal, drug abuse, alcohol abuse, employment, family, and psychiatric. Reliability and validity data for the ASI have been extensively reported (McLellan et al. 1980; McLellan et al. 1985; Rounsaville et al. 1986), and it has become one of the most commonly used evaluation instruments in clinical studies of substance abuse.

Drug History

The drug history was devised for this study. It is a structured interview designed to describe past and current cocaine use in some detail. It also includes questions about age at first cigarette use, level of current cigarette smoking, first illicit drug used, age at which that drug was used, and other

drugs used both at the beginning of cocaine use and during the period of heaviest cocaine use.

FINDINGS

Demographic Characteristics of Clients

The majority of the subjects were male (87 percent) and nonwhite (63 percent black and 21 percent Hispanic). All were between the ages of 20 and 47 (mean age 30.5). Sixty-six percent had 12 years of education or less. Slightly more than three-quarters (77 percent) had held full-time jobs in the 3 years prior to entry into treatment; 20 percent were in professional, executive, or administrative jobs, 11 percent in clerical jobs, and the remaining 69 percent in blue-collar jobs. More than half (57 percent) had been arrested, with about two-thirds (64 percent) on non-drug-related charges.

Drug Use History

All sample members were diagnosed as having both lifetime and current cocaine abuse disorders. Their history of cocaine use ranged from 1 to 32 years, with an average use of 9 years. The mean age at first cocaine use was 21.5, with a range of 10 to 38 years. (Mean age at first use of any illicit drug was 16.2 years, with a range of 8 to 33.) Almost all (84 percent) had used cocaine three or more times a week for at least 1 month, with a third (32 percent) having used it three or more times a week for 4 years or more. Although 96 percent began using cocaine by the intranasal route (snorting), crack or freebase inhalation was the current ingestion route for about three-quarters of the sample (72 percent) in the 30 days prior to first program contact. Half of the sample (48 percent) spent $401 (for at least 1 gram, based on current average prices) on cocaine/crack in the 30 days before first program contact, while 52 percent spent $400 or less.

As expected, most clients were users of cigarettes (89 percent were current smokers), alcohol, and marijuana as well as cocaine. Twenty-three percent had drunk alcohol to intoxication at least three times a week for a month or more at some time in their lives. Seventy-four percent had used marijuana at least three times a week for a period of at least 5 months. Regular use of other drugs was lower—14 percent had used hallucinogens, 11 percent had used amphetamines, and 10 percent or less had used any of a number of other illicit drugs (heroin, other opiates, sedatives, barbiturates, PCP, or Quaalude) three or more times a week for at least 1 month.

Psychopathology of Clients

Of the sample, 47 percent were found to be clinically depressed, according to SCID criteria. This rate is comparable to those reported in other studies of cocaine abusers (Carroll 1989; Gawin and Kleber 1986; Weiss et al. 1986) and opiate addicts (Rounsaville et al. 1982; Woody et al. 1985). Phobic disorders were the only other Axis I diagnoses found in addition to depression, and all persons who were found to have phobic disorders also were diagnosed as having some form of depressive disorder. The four most common Axis II diagnoses were passive-aggressive (20 percent), conduct disorder (19 percent), antisocial personality (18 percent), and paranoid (18 percent). Two studies of opiate addicts have found that at least 40 percent were diagnosed as antisocial personalities (Rounsaville and Weissman 1982; Woody et al. 1985). In two previous studies of cocaine abusers, however, the majority of patients diagnosed as having Axis II disorders were found to be borderline or narcissistically impaired, and no patients were diagnosed as having antisocial personalities (Resnick and Resnick 1986; Weiss et al. 1986). These differences in antisocial personality disorder (ASP) may reflect social class, as the sample of Weiss and coworkers (1986) was predominantly upper middle class (socioeconomic status was not reported for the Resnick and Resnick sample).

The authors attempted to examine the relationship between psychopathology and early use of cigarettes, marijuana, and cocaine by classifying clients according to whether they had: (1) no disorder except substance abuse, (2) Axis II diagnoses but no diagnosis of major depression, or (3) a diagnosis of current or past major depression. Persons with histories of major depression were found to have started using cigarettes, marijuana, and cocaine at significantly earlier ages than had persons in the other two categories (table 1). It is of related interest that recent epidemiological evidence shows that persons with early onset of anxiety or depressive disorders are at increased risk for drug or alcohol dependence (Christie et al. 1988).

Retention in Treatment

Of the subjects, 42 percent (62) were seen for one or two evaluation interviews only and did not return for therapy. Although high, this dropout rate is within the range reported by Baekeland and Lundwall (1975). In addition, 40 percent of all entering clients were seen for three or more therapy sessions, and 25 percent of all entering clients were retained for six or more sessions (figure 1). A distinction was made between early dropout (defined below) and dropout after the first therapy session, and predictors of both early dropout and longer term retention were studied.

TABLE 1. *Mean age at first use of cigarettes, marijuana, and cocaine by psychopathology classification*

Categories	Cigarettes*	Marijuana*	Cocaine*
No Disorder Except Substance Abuse (n=23)	16.7	16.7	23.1
Axis II Disorders Only (n=20)	16.2	15.4	22.0
Axis I Disorders (n=32)	14.1	15.1	19.8

*Probability<.05

NOTE: Data presented in this table are based on a subsample of the total population, consisting of the first 75 people for whom it was possible to administer the SCID.

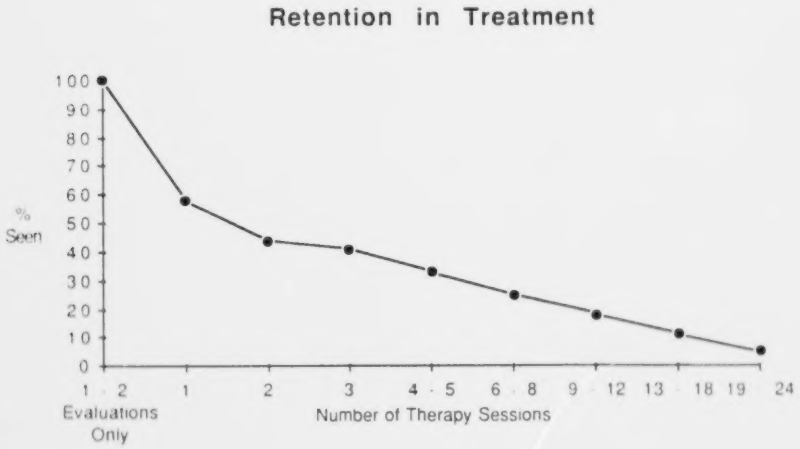

FIGURE 1.

Correlates of Early Dropout. Logistic regression analysis of early dropout (defined as dropping out of treatment before the first therapy session) showed that ethnicity, age, education, number of arrests, global score on the SCL-90, intense lifetime use of marijuana,[1] intense lifetime use of cocaine,[2] and treatment assignment were not significantly associated with early dropout. The strongest predictor appeared to be therapist assignment. The rate of retention varied radically among the seven SE and family therapists. Only 14 percent of the clients of one therapist (n=7) were retained for four or more sessions, while 81 percent of the clients of another (n=16) were retained for that long. The range for the remaining five therapists fell between 40 percent and 69 percent.

Predictors of Longer Term Retention. Multiple regression analysis showed that ethnicity, SCL-90 score, arrest history, age, intense lifetime use of marijuana, intense lifetime use of cocaine, type of therapy, and education accounted for 25 percent of the variance in retention for one or more therapy sessions (table 2). Whites and younger clients were significantly

TABLE 2. *Regression of retention in therapy on predictor variables*

	Retention in Therapy (n=86)
Variable	Beta
White (Contrasted With Black)	.26*
Age	-.24*
No Arrest (Contrasted With One or More Arrests)	-.21
Hispanic (Contrasted With Black)	.21
SCL—Total Score	-.19
Intense Lifetime Use of Marijuana[1]	.17
Intense Lifetime Use of Cocaine[2]	-.16
Individual Therapy	.11
Family Therapy	.07
Education	.03
Multiple R^2	.25
Overall F	2.42**

*p≤.05

**p≤.01

[1]Defined as use of marijuana three or more times a week for a period of at least 1 month at some time in life.

[2]Defined as use of cocaine three or more times a week for a period of at least 1 month at some time in life.

more likely to be retained in treatment for a relatively long period of time. There was a trend indicating that number of arrests and level of symptomatic distress were both negatively related to retention. Although both individual and family therapy were associated with a trend toward longer retention than was treatment as usual, the relationship was not significant. As mentioned earlier, these retention rates by treatment were highly interrelated with therapist assignment.

Cocaine Use During Treatment

Cocaine use during the treatment period was determined by two independent measures: urinalysis and therapist report. Therapist report of cocaine use was recorded after each session, while a urine sample was taken after every third session. A client was considered to be using cocaine if the therapist reported it or if one or more urine tests were positive for cocaine. Hand-tabulation of results for the subsample of 55 clients who were retained in treatment for three or more sessions and for whom urine tests and therapist reports were available showed that 25 percent (14) did not use cocaine at any time during treatment.

Cocaine Use 6 Months After Entry Into Treatment

Similarly, there were two sources of information about cocaine use 6 months after entry into treatment: analysis of the urine sample submitted at time of followup and the client's self-report about cocaine use throughout the 6-month period and in the 30 days immediately prior to the followup interview. Preliminary analysis of findings for the first 70 clients for whom both self-report and urinalysis results were available indicated that 27 percent (19) were free of cocaine use, according to both urine test and self-report, in the followup period. Further analysis of data, including relationship of outcome to form of psychotherapy offered, will be reported at a later time.

DISCUSSION

These data show that therapist assignment was the strongest predictor of who will remain in treatment. This is similar to recent work by Luborsky and coworkers (unpublished manuscript), who found significant differences in retention and outcome between therapists in a range of psychotherapy studies and suggest that therapist factors can be strong predictors of retention and outcome.

The preliminary findings on cocaine use during and after treatment suggest that roughly one-quarter of these clients ceased use of cocaine during the period under study. The fact that a substantial minority of cocaine addicts stopped their use in the postcontact period is noteworthy because many patients with this problem are routinely assigned to expensive 28-day

inpatient treatment. For some of our clients, cessation of use may be an outcome of treatment, while for others it may represent "spontaneous remission." Subsequent analysis will explore this issue further and will also determine the characteristics that distinguish remitters from others.

It is possible that some of the treatment failures of this study may prove to be the successes of future treatment attempts. The therapy provided as part of this study was the first treatment experience for 84 percent of the population studied here, and research based on populations of heroin addicts indicates that those who have had two or more treatment attempts are more likely to have favorable outcomes (Brown et al. 1972; Kleinman and Lukoff 1980; McLellan and Druley 1977; Williams and Johnston 1972).

The overall conclusion at this point of the data analyses is that low intensity psychotherapy or family therapy, as offered to the clients in this study, is insufficient to produce remission of cocaine use in the majority of patients. More frequent contact on an outpatient basis, such as that now being attempted in a number of programs serving cocaine addicts (Millman 1989; Todd 1989; Woody 1989), may prove to be more effective, as may residential programs providing 30 days or more of treatment. Regular attendance at self-help groups, such as Cocaine Anonymous, also may prove useful to persons who are willing to maintain involvement with them, though few data are available in this area.

FOOTNOTES:

1. Defined as use of marijuana three or more times a week for a period of at least 1 month at some time in life.

2. Defined as use of marijuana three or more times a week for a period of at least 1 month at some time in life.

REFERENCES

Baekeland, F., and Lundwall, L. Dropping out of treatment: A critical review. *Psychol Bull* 82:738-783, 1975.

Bateson, G.; Jackson, D.; Haley, J; and Weakland, J. Toward a theory of schizophrenia. *Behav Sci* 1(4):251-264, 1956.

Beck, A., and Steer, R. *Beck Depression Inventory Manual.* New York: Harcourt Brace Jovanovich, Inc., 1981.

Brown, B.S.; Du Pont, R.L; and Glendinning, S.T. Narcotics treatment and behavioral change. In: *Proceedings of the Fourth National Conference on Methadone Treatment.* New York: National Association for the Prevention of Addiction to Narcotics, 1972. pp. 163-168.

Carroll, K. Psychotherapy and counseling in the treatment of cocaine abuse. Presented at the National Institute on Drug Abuse Technical Review Meeting on Psychotherapy and Counseling in the Treatment of Drug Abuse, Rockville, MD, May 18, 1989.

Christie, K.A.; Burke, J.D.; Regier, D.A.; et al. Epidemiologic evidence for early onset of mental disorders and higher risk of drug abuse in young adults. *Am J Psychiatry* 145(8):971-975, 1988.

Derogatis, L. *SCL-90-R: Administration, Scoring and Procedures Manual-II.* Towson, MD: Clinical Psychometrics Research, 1983.

Gawin, F.H., and Kleber, H.D. Abstinence symptomatology and psychiatric diagnosis in cocaine abusers: Clinical observations. *Arch Gen Psychiatry* 43:107-113, 1986.

Gibbon, M., personal communication, 1987.

Haley, J. *Uncommon Therapy.* New York: Norton, 1973.

Haley, J. *Leaving Home.* New York: McGraw Hill, 1980.

Hoffman, L. *Foundations of Family Therapy.* New York: Basic Books, 1981.

Jackson, D. The question of family homeostasis. *Psychiatr Q* 31:79-90, 1957.

Kleber, H.D., and Gawin, F.H. The spectrum of cocaine abuse and its treatment. *J Clin Psychiatry* 45:18-23, 1984.

Kleinman, P.H., and Lukoff, I.F. *Methadone maintenance: Modest help for a few.* Final report to the National Institute on Drug Abuse of the research project on evaluation of the effects of methadone, 1975.

Kleinman, P.H., and Lukoff, I.F. The Life Cycle of Treated Heroin Addicts. Report of the project on the life cycle of treated heroin addicts to the National Institute on Drug Abuse. Unpublished manuscript, 1980.

Luborsky, L.; Woody, G.E.; Hole, A.; and Velleco, A. Treatment Manual for Supportive-Expressive Psychoanalytically Oriented Psychotherapy. Unpublished manuscript, 1977.

McLellan, A., and Druley, K. The re-admitted drug patient. *Hosp Community Psychiatry* 28:764-766, 1977.

McLellan, A.T.; Luborsky, L.; Cacciola, J.; Griffith, J.; Evans, F.; Barr, H.L.; and O'Brien, C.P. New data from the addiction severity index: Reliability and validity in three centers. *J Nerv Ment Dis* 173:412-428, 1985.

McLellan, A.T.; Luborsky, L.; Woody, G.E.; and O'Brien, C.P. An improved diagnostic evaluation instrument for substance abuse patients: The addiction severity index. *J Nerv Ment Dis* 168:26-33, 1980.

Millman, R.B., personal communication, 1989.

Olson, D.H.; Portner, J.; and Lavee, Y. *FACES III.* St. Paul, MN: Family Social Science, University of Minnesota, 1985.

Resnick, R.B., and Resnick E. Psychological issues in the treatment of cocaine abuse. In: Harris, L.S., ed. *Problems of Drug Dependence, 1985.* National Institute on Drug Abuse Research Monograph 67. DHHS Pub. No. (ADM) 86-1448 Rockville, MD: the Institute, 1986. pp. 291-294.

Rosecan, J.S.; Spitz, H.I.; and Gross, B.G. Contemporary issues in the treatment of cocaine abuse. In: Spitz, H.I., and Rosecan, J.S., eds. *Cocaine Abuse: New Directions in Treatment and Research.* New York: Brunner Mazel, 1987.

Rounsaville, B.; Kosten, T.; Weissman, M.M.; and Kleber, H.D. Prognostic significance of psychopathology in treated opiate addicts. *Arch Gen Psychiatry* 43:739-745, 1986.

Rounsaville, B., and Weissman, M. Diagnosis and symptoms of depression in opiate addicts. *Arch Gen Psychiatry* 39:151-156, 1982.

Rounsaville, B.; Weissman, M.; Kleber, H.; and Wilbur, C. Heterogeneity of psychiatric diagnosis in treated opiate addicts. *Arch Gen Psychiatry* 39:161-166, 1982.

Schachter, S. Recidivism and self-cure of smoking and obesity. *Am Psychol* 37(4):436-444, 1982.

Schuster, C. Initiatives at the National Institute on Drug Abuse. Problems of drug dependence 1987: Proceedings of the 49th annual scientific meeting. National Institute on Drug Abuse Research Monograph 81. DHHS Pub. No. (ADM)88-1564. Rockville, MD: the Institute, 1988.

Spitzer, R.; Williams, J.; and Gibbon, M. *Instruction manual for the structured clinical interview for DSM-III-R (SCID 4/1/87 revision).* New York: New York State Psychiatric Institute, 1987.

Stanton, M.D.; Todd, T.C.; et al. *The Family Therapy of Drug Abuse and Addiction.* New York: The Guilford Press, 1982.

Todd, T.C., personal communication, 1989.

Todd, T.C.; Stanton, M.D.; and Carway, J. A Manual for Structural Strategic Family Therapy. Unpublished manuscript, 1985.

Washton, A.M., and Gold, M.S. *Cocaine: A Clinician's Handbook.* New York: Guilford Press, 1987.

Weiss, R.D.; Mirin, S.M.; Michael, J.L.; and Sollogub, A. Psychopathology in chronic cocaine abusers. *Am J Drug Alcohol Abuse* 12:17-29, 1986.

Williams, H.R., and Johnston, B.A. Factors related to treatment retention in a methadone maintenance program. In Proceedings of the Fourth National Conference on Methadone Treatment. New York: National Association for the Prevention of Addiction to Narcotics, 1972. pp. 439-442.

Woody, G.E., personal communication, 1989.

Woody, G.E.; Luborsky, L.; McLellan, A.T.; O'Brien, C.P.; Beck, A.T.; Blaine, J.; Herman, I.; and Hole, A. Psychotherapy for opiate addicts: Does it help? *Arch Gen Psychiatry* 40:639-645, 1983.

Woody, G.E.; McLellan, A.T.; Luborsky, L.; and O'Brien, C.P. Sociopathy and psychotherapy outcome. *Arch Gen Psychiatry* 42:1081-1086, 1985.

AUTHORS

Paula H. Kleinman, Ph.D.
George E. Woody, M.D.
Thomas C. Todd, Ph.D.
Robert B. Millman, M.D.
Sung-Yeon Kang, Ph.D.
Jack Kemp, M.S.
Douglas S. Lipton, Ph.D.

Narcotic and Drug Research, Inc.
11 Beach Street
New York, NY 10013

Methodological Considerations

Therapist Effects in the Treatment of Drug Dependence: Implications for Conducting Comparative Treatment Studies

Paul Crits-Christoph, Katherine L. Beebe, and Mary Beth Connolly

INTRODUCTION

An overlooked issue in the design of psychotherapy outcome studies is the possibility of systematic differences between therapists. Crits-Christoph and Mintz (in press) have recently reported on the implications of therapist differences for the design and analysis of comparative psychotherapy outcome studies. The purpose of this chapter is to bring this problem of "therapist effects" to the attention of researchers in the area of drug-dependence treatment. The chapter has four aims: (1) to describe the statistical implications of differences between therapists for testing treatment effects; (2) to present the results of a literature review summarizing how investigators have analyzed their data in terms of therapist differences; (3) to report on the size of therapist effects in eight major psychotherapy outcome studies (including one involving drug-dependent patients); and (4) to illustrate the implications of ignoring therapist effects, using artificial data to estimate the amount of bias in significance testing when the incorrect analysis is performed.

STATISTICAL IMPLICATIONS OF THERAPIST EFFECTS

Consider a typical psychotherapy outcome study: Two or three treatment conditions are tested against each other in terms of efficacy. Usually, each treatment is performed by a few trained experts in the particular treatment modality. In analysis of variance language, these therapists may be "nested" within treatments (different therapists perform each treatment) or crossed with treatments (the same therapists conduct each form of treatment). How should these designs be analyzed to assess for treatment effects?

The statistical implications of therapist effects for testing treatment effects have been described by Martindale (1978). In brief, the statistical issue involved here is the problem of violation of the assumption of non-independence in the analysis of variance. The serious consequences of violations of this assumption have recently been described by Kenny and Judd (1986). In regard to psychotherapy outcome research, the question reduces to the following: (1) If therapists are, like patients, sampled for a particular study; (2) if the researcher wishes to generalize results beyond the particular sample of therapists used in the study; and (3) if there are systematic differences between therapists on the dependent variable, then a therapist factor must be included as a "random" factor, rather than ignored or treated as a "fixed" factor, in the ANOVA design. If there are systematic differences between therapists, and therapist is not specified as a random factor, we have violated the important ANOVA assumption of non-independence of observations, and the generalization of the results is questionable.

The distinctions between "random" and "fixed" factors are described in any basic ANOVA text (Winer 1971). A factor is random when its levels are drawn at random from a pool of possible levels. With a fixed factor, all possible levels of interest are included within the study. The use of a random factor allows the experimenter to generalize the results of a study to levels of the variable not used, i.e., the pool, whereas designating a factor as fixed means the experimenter does not wish to generalize beyond the levels used in the study. For example, subjects are generally a random factor in an experiment: We wish to generalize the results of the study to a larger pool of potential subjects. A comparison of the effects of two drug conditions, however, is generally a fixed factor. In this case, we do not wish to generalize to other drugs, but rather conclusions are limited to the drugs tested.

We argue that in studies of the comparative efficacy of psychotherapies, therapists should logically be treated as a random factor; that is, we wish to generalize the results of the study to the pool of similarly trained and selected therapists. This allows us, for example, to make a conclusion such as cognitive/behavioral psychotherapy is superior to psychodynamic psychotherapy for the treatment of cocaine abuse. A conclusion in which we do not generalize to other therapists would read as follows: Cognitive/behavioral psychotherapy as practiced by Tim, Bob, and Fred is superior to psychodynamic psychotherapy as practiced by David, Anita, and Jacques. The results may or may not be relevant to other therapists. Such a conclusion has limited scientific value. Other researchers and clinicians could not make use of these results unless they hired the same therapists to perform treatment. One other point should be noted: Specifying therapist as a random factor does not imply that all therapists are interchangeable and therefore the same (Paul and Licht 1978). On the contrary, because therapists are different and potentially have a different impact on patients, they

must be considered a random factor so that we can take these differences into account in our conclusions.

How does designating therapist as a random factor change our tests of significance of treatment effects? An example will serve to illustrate the consequences of treating therapists as a random factor. Let us say we are comparing two treatments, with each treatment performed by the same number of therapists and therapists nested within treatments. If we ignore the therapist factor, a comparison of two treatments is a simple one-way analysis of variance, with the F test for treatment differences equal to mean square between treatment groups divided by mean square error (within cell). The degrees of freedom for this F test are 1 for the numerator and the number of subjects minus 2 for the denominator. Now, if "therapist" is designated as a random factor, the significance test for treatment effects becomes mean square between treatment groups divided by mean square for therapists. The degrees of freedom are now 1 and the number of therapists minus 2. Thus, the significance test has changed in two ways. First, treatment effects are judged relative to therapist effects, not differences between subjects (within cell error). If therapist differences are not trivial, the denominator becomes larger, making it more difficult to obtain a high F ratio. The second change is in the degrees of freedom. The effective sample size for significance testing with therapist as a random effect is related to the number of therapists, not to the number of subjects. Since the number of therapists is generally less than the number of subjects in a study, we have much lower statistical power for detecting a significant treatment effect.

A more thorough discussion of the issues involved, including what happens when therapists are treated as a fixed effect and designs where therapists and treatments are crossed, is provided by Martindale (1978). The main point we wish to make is that significance testing for treatment effects is changed depending on whether or not the therapist factor is specified as random.

HOW HAVE RESEARCHERS TREATED THE THERAPIST FACTOR?

Martindale (1978) reviewed treatment-outcome studies published in the 1975 issues of the *Journal of Consulting and Clinical Psychology* and the 1973 and 1974 issues of the *Journal of Abnormal Psychology*. In only one of 33 articles was the therapist factor correctly specified as a random factor.

We were interested in the extent to which researchers have become aware of this issue since the 1978 Martindale article. All treatment studies published in the *Journal of Consulting and Clinical Psychology* from 1979 until 1986 were examined. As was found in Martindale's (1978) review, the therapist factor was often completely ignored (56 of the 108 studies).

For 26 of the 108 studies, there was only one therapist in the study or one therapist per treatment, making it impossible to use therapist as a separate factor.

In 23 studies, therapist was treated as a fixed factor, or a preliminary one-way analysis of variance was performed to rule out therapist effects. In these studies, however, the p value for deciding that there is no evidence of a therapist effect was set at $p>.05$ or was not specified. As Martindale (1978) points out, setting the p value for ruling out therapist effects at .05 is inappropriate because the test for treatment effects will still be significantly affected even if the test for therapist differences does not reach the .05 level. Kirk (1968) has recommended that p be set at .25, and Winer (1971) suggests .20 or .30 for ruling out a factor in a preliminary analysis.

Of the 108 studies, the therapist factor was correctly analyzed in only three. This meant that either the therapist was treated as a random factor or therapist effects were ruled out in a preliminary analysis with a $p>.20$. Thus, psychotherapy researchers in general have continued to ignore Martindale's (1978) advice about a straightforward analysis-of-variance issue.

THERAPIST EFFECTS IN EIGHT TREATMENT-OUTCOME STUDIES

Perhaps psychotherapy researchers have chosen to ignore the therapist factor because they believe that therapist differences in outcome generally do not exist. If therapist differences are zero, then the therapist factor can be ignored and the F ratio for treatment effects comparing mean square for treatment to mean square error (within cell) is appropriate.

We obtained the raw data from several completed psychotherapy outcome studies in order to determine whether therapist effects are typically zero. The studies obtained were: (1) Beck et al. 1985 and Rush et al. 1977 (combined data from these two studies), (2) Zitrin et al. 1978, (3) Thompson et al. 1987, (4) Pilkonis et al. 1984, (5) Piper et al. 1984, (6) Nash et al. 1965, (7) Hollon et al., unpublished manuscript, and (8) Woody et al., unpublished manuscript.

The data from the combined studies of Beck and coworkers (1985) and Rush and coworkers (1977) consisted of depressed patients treated with either cognitive therapy alone or cognitive therapy plus amitriptyline hydrochloride. For our analysis, we used only those therapists who had at least three patients. This left a total of 6 therapists who treated a total of 30 patients. In study 2, 5 therapists treated 71 patients with behavior therapy plus imipramine, and 5 therapists treated 79 patients with supportive therapy plus imipramine. In study 3, 3 therapists treated 35 depressed patients with behavior therapy, 5 therapists treated 44 patients with cognitive

therapy, and 3 therapists treated 35 patients with dynamic therapy. In study 4, 3 therapists treated 22 mixed-diagnosis outpatients with psychodynamic therapy. In study 5, patients received short-term individual therapy (n=21), long-term individual therapy (n=20), short-term group therapy (n=19), or long-term group therapy (n=19). Three therapists carried out psychodynamically oriented treatment, with each therapist treating between five and eight patients in each therapy modality. In study 6, 4 therapists treated 40 neurotic outpatients with dynamic therapy. In study 7, 4 therapists treated 16 depressed patients with cognitive therapy alone, and the same therapists treated 16 patients with cognitive therapy plus imipramine. In study 8, 3 therapists treated 30 opiate-addicted patients with dynamic therapy.

For each study, analyses of variance were performed on each outcome measure, employing treatment group and therapist as factors. The therapist factor was specified as a random term. Percent of variance due to therapist was calculated from the equations for expected mean squares.

In order to obtain an overall sense of the size of therapist effects within each study, we averaged the percentages of outcome variance due to therapist across all of the measures used within each study. The number of outcome measures used within each study varied from 1 to 18 (median=7.5).

The results for these average effect sizes for each study are presented in table 1. Notable is the large variability in therapist effects across studies. The combined data of Beck and coworkers (1985) and Rush and coworkers (1977) and the study of Thompson and coworkers (1987) both yielded therapist effects equal to zero. Most of the other studies had modest to average therapist effects (in the 5 percent to 10 percent of variance range). One study (Nash et al. 1965) had an average therapist effect across six outcome measures equal to 29 percent—a very large effect.

The average effect for each study may obscure potentially large effects on certain measures within a study. Table 1 presents results for the largest therapist effect found within each study. As can be seen, many of the studies evidenced large effects on at least one of the outcome measures used. For example, in the study of Pilkonis and coworkers (1984), 23 percent of the outcome variance was due to therapist on one of the measures. We found a 43-percent effect on one of the measures in the study of Piper and coworkers (1984). For Nash and coworkers (1965), a 55-percent effect was found, and for Woody and coworkers (unpublished manuscript), a 20-percent effect was found. These kinds of effects would generally be seen as large effects for the behavioral sciences (Cohen 1969).

The reasons for the striking variability in therapist effects across studies and across measures within studies are unknown. We can speculate that different forms of treatment may be more prone to therapist effects; for example,

unstructured treatments like psychodynamic therapy may give more leeway for aspects of the therapist to come into play, but structured treatments, such as systematic desensitization, might be less subject to therapist effects.

TABLE 1. *Therapist effects in eight studies*

	Percent of Outcome Variance Due to Therapist	
Study	Average Across Measures (%)	Largest Effect (%)
Woody et al. (unpublished manuscript)	4.3	20.1
Pilkonis et al. (1984)	7.6	23.4
Nash et al. (1965)	28.7	54.7
Piper et al. (1984)	7.0	43.1
Zitrin et al. (1978)	6.1	15.6
Thompson et al. (1987)	0	0
Beck et al. (1985)/ Rush et al. (1977)	0	0
Hollon et al. (unpublished manuscript)	5.2	5.2

Additionally, the processes used to select, train, and supervise therapists in a given study may be responsible for differences in the range in quality of therapists within a study. We are conducting further research on these questions to determine whether we can understand the reasons for differences in therapist effects.

BIAS IN SIGNIFICANCE TESTING

We have presented the statistical argument for considering therapist to be a random term and have discussed the changes in significance testing that follow when the appropriate analysis is done. A review of the literature has indicated that in most cases, the appropriate analysis is not performed. We have also provided evidence that therapist effects can, at times, be large in psychotherapy studies. A final question remains: How much of a bias in significance testing is introduced when therapist effects are large but the typical inappropriate analysis is conducted? In other words, to what extent are the conclusions reported in such studies incorrect?

The extent to which the significance testing in the analysis of variance is distorted by not including a factor as random when it should be has not, to our knowledge, been specified in the literature on analysis of variance. Kenny and Judd (1986) give formulas for calculating the effects of nonindependence on expected mean squares. Nonindependence, however, also affects the F ratio by producing more variable estimates of the mean squares and by producing a correlation between mean square treatment and mean square error (Kenny and Judd 1986). The effects of all of these sources of distortion were recently investigated by Crits-Christoph and Mintz (in press).

Stated simply, the issue is this: How many times is a false-positive treatment effect obtained when the incorrect analysis is performed, i.e., when the therapist factor is ignored? Crits-Christoph and Mintz (in press) approached this problem by conducting a series of simulated psychotherapy outcome studies, and their methods and results will be briefly described here. Artificial data were created via computer random-number procedures to analyze a variety of ANOVA designs. The basic design examined included three treatment groups, with therapists nested within treatments. The authors then varied the number of therapists per treatment (2, 5, 10, and 15 therapists were tried) and the number of patients per therapist (2, 4, 8, and 15 therapists per patient were tried). Finally, built into the "outcome scores" were various effects due to therapist (5, 15, and 25 percent due to therapist were tried). Note that by building in a systematic effect due to therapist, the authors did not change the fact that no systematic treatment effects should exist, since the scores are still random numbers in regard to treatment group. Once a particular design was specified, e.g., two therapists per treatment, four patients per therapist, or 15 percent of the variance due to therapist, Crits-Christoph and Mintz (in press) performed an ANOVA, ignoring the therapist factor, and tabulated whether or not a significant (at $p<.05$) treatment effect was found. This procedure was done 2,000 times on each design, using new random numbers with each trial. Since no treatment effect existed, it would be expected that only 5 percent of the trials to yield a p value of .05 or less. The actual percent of "significant" findings over the 2,000 trials represents the probability of occurrence of false-positive treatment effects. Note that these calculations are not specific to psychotherapy but rather apply to any experiment with the types of designs given above.

The results indicated that the number of treatments and the number of therapists per treatment did not affect the probability of false-positive treatment effects to any extent. The number of patients per therapist and the percent of variance due to therapist, however, showed systematic relationships to the probability of false-positive treatment effects. The probability of false-positive treatment effects rises linearly with the number of patients per therapist, and this increase is more pronounced for higher percentages due to therapist. At the extreme, a design with 15 patients per

therapist and 25 percent of the variance due to therapist yields a probability of false-positive treatment effects of .52. A design of eight patients per therapist and 15 percent of the outcome variance due to therapist yields a false-positive probability of .23. These false-positive rates are clearly unacceptable.

IMPLICATIONS

There are several important ramifications of these findings. For one, the results of some published studies are in question. Studies in the literature with modest to large therapist effects may have reported misleading conclusions; that is, the presence of therapist effects may have led to conclusions that treatments differ when in fact they do not.

These results may also have implications for other types of treatment studies besides psychotherapy outcome trials. For example, to the extent that nonspecific factors in the doctor/patient relationship are important, "doctor" effects may be present in trials involving pharmacotherapy. The use of double-blind designs and the same physicians in each treatment condition, however, may effectively control for the problem. Although we have focused on psychotherapy outcome, the problem of therapist effects is also germane to studies of the process of psychotherapy as well.

Treatment research on substance abuse raises some related issues. To the extent that drug counselors produce differential outcomes, a "counselor" factor should also be included in designs as a random factor. Recent data reported by McLellan and coworkers (1988) suggest that different drug counselors may indeed have different rates of success. To complicate things even further, substance-abuse treatment is usually administered within special treatment facilities. These facilities or programs may be an additional factor over which we wish to generalize our results. The relevant research question is as follows: Can another investigator replicate our results if a study is done in similar treatment programs using similarly trained and selected therapists and counselors? Including program, therapist, and counselor as random factors in the study design allows for this question to be answered. It may not be practical, however, to include these factors in a design and to obtain enough degrees of freedom for significance testing with adequate statistical power. Other possible solutions are given below.

A final implication of the results presented here is that therapist effects are not just a nuisance that must be accounted for when examining treatment differences. Rather, differences between therapists may be important data in their own right. Studies of particularly effective vs. noneffective therapists may be a useful way to pursue an understanding of how psychotherapy works and how we can better train therapists to be successful.

RECOMMENDATIONS

We can offer several recommendations to researchers conducting studies of psychotherapy with substance-abuse patients. Foremost, investigators should routinely test for therapist differences. A lack of information about therapist differences not only leaves the study vulnerable to potentially erroneous conclusions but prevents researchers from understanding a potentially interesting and meaningful source of outcome variation.

If an investigator chooses to ignore therapist effects, perhaps because statistical power would be too low if therapist was included as a factor, our recommendation is that the treatment-outcome data be interpreted cautiously. This means including a sentence in the discussion section of the research report stating that the results may not hold up if other similarly trained therapists are used in a replication attempt.

In terms of planning a treatment-outcome study, our recommendation is that as many therapists as possible be included in the study if there is reason to believe that therapist differences may be evident. By maximizing the number of therapists, the researcher will have the most degrees of freedom for testing treatment effects. On the other hand, the investigator may wish to avoid the whole issue by trying to insure that no therapist effects are present. By careful selection, training, and supervision of therapists, it may be possible to reduce therapist variability. A clearly specified treatment approach given in a manual is also likely to minimize differences among therapists. The extent to which these procedures can insure a lack of therapist differences is, however, an unknown question at this point. Thus, we advocate employing these procedures but also checking on the extent to which therapists are still producing differential outcomes.

REFERENCES

Beck, A.T.; Hollon, S.D.; Young, J.E.; Bedrosian, R.C.; and Budenz, D. Treatment of depression with cognitive therapy and amitriptyline. *Arch Gen Psychiatry* 42:142-148, 1985.

Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press, 1969.

Crits-Christoph, P., and Mintz, J. The implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol*, in press.

Hollon, S.D.; DeRubeis, R.J.; Evans, M.D.; Wiemer, M.J.; Garvey, M.J.; Grove, W.M.; and Tuason, V.B. Cognitive therapy, pharmacotherapy, and combined cognitive-pharmacotherapy in the treatment of depression: I. Differential outcome in the CPT project. Unpublished manuscript.

Kenny, D.A., and Judd, C.M. Consequences of violating the independence assumption in analysis of variance. *Psychol Bull* 99(3):422-431, 1986.

Kirk, R.E. *Experimental Design Procedures for the Behavioral Sciences.* Belmont, CA: Brooks/Cole, 1968.

Martindale, C. The therapist-as-fixed-effect fallacy in psychotherapy research. *J Consult Clin Psychol* 46(6):1526-1530, 1978.

McLellan, A.T.; Woody, G.E.; Luborsky, L.; and Goehn, L. Is the counselor an "active ingredient" in substance abuse rehabilitation? An examination of treatment success among four counselors. *J Nerv Ment Dis* 176(7):423-430, 1988.

Nash, E.; Hoehn-Sacric, R.; Battle, C.; Stone, A.; Imber, S.; and Frank, J. Systematic preparation of patients for short-term psychotherapy—II: Relation to characteristics of patient, therapist and the psychotherapeutic process. *J Nerv Ment Dis* 140:374-383, 1965.

Paul, G.L., and Licht, M.H. Resurrection of uniformity assumption myths and the fallacy of statistical absolutes in psychotherapy research. *J Consult Clin Psychol* 46(6):1531-1534, 1978.

Pilkonis, P.A.; Imber, S.D.; Lewis, P.; and Rubinsky, P. A comparative outcome study of individual, group, and conjoint psychotherapy. *Arch Gen Psychiatry* 41:431-437, 1984.

Piper, W.E.; Debbane, E.G.; Bienvenu, J.P.; and Garant, J. A comparative study of four forms of psychotherapy. *J Consult Clin Psychol* 52(2):268-279, 1984.

Rush, A.J.; Beck, A.T.; Kovacs, M.; and Hollon, S.D. Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Ther Res* 1:17-37, 1977.

Thompson, L.W.; Gallagher, D.; and Breckenridge, J.S. Comparative effectiveness of psychotherapies for depressed elders. *J Consult Clin Psychol* 55(3):385-390, 1987.

Winer, B.J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill, 1971.

Woody, G.; McLellan, A.T.; O'Brien, C.; Woorsky, L. Supportive expressive psychotherapy in the treatment of opiate dependence. Unpublished manuscript.

Zitrin, C.M.; Klein, D.F.; and Woerner, M.G. Behavior therapy, supportive psychotherapy, imipramine, and phobias. *Arch Gen Psychiatry* 35:307-316, 1978.

ACKNOWLEDGMENT

Preparation of this article was supported by the National Institute of Drug Abuse grant DA 04840-03.

AUTHORS

Paul Crits-Christoph, Ph.D.
Katherine L. Beebe, B.A.
Mary Beth Connolly, B.A.

Department of Psychiatry
University of Pennsylvania
Philadelphia, PA 19104

Control Groups and Comparison Groups in Psychotherapy Outcome Research

T.D. Borkovec

INTRODUCTION

The customary goals of a therapy outcome design are to demonstrate the efficacy of a therapy procedure and to identify the mechanisms of the procedure by which it produces change. The former has applied significance, while the latter adds to our knowledge about the nature of human behavior from which to derive further applied methods and progress in understanding. As is the case with any scientific approach to a question, both of these goals are most efficiently accomplished by the application of strong inference (Platt 1964), the systematic ruling out of rival hypotheses regarding the cause-and-effect relationships inherent in observed relationships, taking what remains unrejected, and creating and testing further rival hypotheses about what is left over.

Any observed outcome of a therapy procedure, irrespective of the design employed, potentially bears on both of the goals, including the case study. But different designs vary in specificity, that is, how many relevant hypotheses about cause can be rejected, and this will vary as a function of the contrast conditions chosen for the design. The ideal design holds everything constant but one element (a "single" element in a relative sense) that is systematically varied across the different conditions employed. Control conditions are used to help in ruling out factors that are not causatively related to change. They do so by containing factors that are common to both their procedures (this is how they hold things constant) and to the therapy condition to which they are compared. Common factors, even though they may be important or essential in interaction with noncommon elements in producing change, cannot by themselves explain any differential outcome between conditions.

TYPES OF COMPARISON/CONTROL DESIGNS

Commonly used designs can be ordered in terms of the degree to which their control conditions allow increasingly specific cause-and-effect conclusions, by virtue of their ability to rule out increasingly numerous rival hypotheses regarding what factors could have caused the observed improvement.

Therapy/No-Treatment Comparison Design

The most elemental experimental design is the therapy/no-treatment comparison. Ordinarily, the no-treatment clients are told that therapy cannot begin for a certain amount of time, that time being matched to the duration of the treatment trial for therapy clients. It is thus a waiting-list no-treatment condition. Clients in no-treatment do, however, undergo the same pretesting and posttesting.

When significant change is observed in clients undergoing therapy, the possibility that therapy caused the change is supported but so also are numerous other possible reasons for the change: history, maturation, repeated assessment, statistical regression, changes in the measurement devices, and the selective attrition of clients who would have otherwise shown no change (Campbell and Stanley 1963). Thus, the design provides a control condition that contains these factors in common with the therapy condition but does not contain the therapy element. A no-treatment condition does this, since by random assignment both conditions have, in the long run (if repeated many times), an equal chance of being affected by these factors. Any observed difference between the therapy and the control condition can be attributed to the way in which they differ (presence or absence of therapy), and ways in which they are similar (presence of history, maturation, etc.) can be ruled out.

This design does not come very close to approximating the design ideal of holding everything but one element constant. Therapy differs from no therapy in many different ways, and thus there are many rival hypotheses for why therapy may produce greater improvement. Remaining design options will help to rule out some of these other, uncontrolled factors, thus allowing more specificity about what caused the change. But the Treatment/No-Treatment Design is useful in the initial stages of a research program. It is often less costly in time and money than more powerful designs, and it may be wise to determine whether a therapy has any potential value at all, relative to no therapy, before proceeding with more sophisticated designs. It needs to be acknowledged, however, that the payoff in terms of knowledge acquisition is minimal.

For ethical reasons, clients assigned to a no-treatment condition are provided with treatment at the end of the posttest period. While this precludes long-

term followup information on untreated clients, it does provide a valuable opportunity to replicate therapy effects (thus, the importance of continuing customary outcome assessments with these clients), and the ethical requirements of providing therapy within a reasonable period of time to participants take precedence over scientific ideal. For this same reason, one needs to be cautious when the study involves an extremely lengthy treatment-trial period. Ethical decisions here will be partly a function of the nature and course of the diagnostic problem being treated. For many cases of outpatient anxiety disorder and depression, for example, therapy trials of 2 to 3 months have been considered reasonable. Finally, it is incumbent on the investigator to have some method for handling crises or deterioration in no-treatment clients. This is often done by telling the clients to contact the investigator if significant problems arise, by maintaining periodic phone contacts to assess the client's status, or by monitoring daily self-report measures during the wait period. If deterioration is identified, the client needs to be removed from the study and provided with appropriate treatment.

Placebo or Nonspecific Comparison Design

From a strong inference point of view, the task at any stage of research on a therapy procedure is to take what is left over after a comparison, i.e., take what is uncommon relative to the control condition and has not yet been ruled out as a cause of change, and break it apart into its more specific elements. The investigator then creates control conditions that represent one or more of these elements. Significant differences among these conditions would allow the ruling out of certain elements and would narrow the experimental search for active ingredients to a more specific aspect of the total therapy package. Thus, after a therapy/no-treatment comparison, one takes what is left over (therapy, in general) and examines what complex set of factors remains represented that may be the cause of change.

One major set of factors contained in therapy is the nonspecific factor. The change that was observed may have been due to factors common to any form of therapeutic intervention and not specific to this particular therapy. These factors include elements such as expectation for improvement, contact with a therapist, suggestion effects, and attention to the problem. Research interest is usually in a particular therapy, and if it is not superior in efficacy to what is common to therapies in general, there is no reason to get excited about its potential usefulness. Thus, the Placebo or Nonspecific Comparison Design became for a while the sine qua non of outcome methodology (Paul 1969). If a therapy can be shown to produce change greater than that induced by a placebo condition, then those factors common to all therapies can be ruled out as the sole explanation of change, leaving unrejected something specific about the therapy. In addition, the placebo condition provides another opportunity to rule out the effects of history, maturation, and so on, since these factors are as equally operative in the placebo group

as they are in the no-treatment condition. When therapy is superior to a placebo condition, evidence exists that something specifically active is contained in the therapy and is causing the change. Upon demonstration of such an effect, there is greater optimism that further research on the therapy is likely to lead to the development of more effective forms of the therapy, the determination of the particular client subtypes that are particularly responsive to this form of treatment, the identification of specific mechanisms by which the therapy produces change, and so on.

In early stages of research on a therapy technique with a particular diagnostic group, a no-treatment condition will ordinarily be included along with a nonspecific control condition. After several investigations that demonstrate little change in client problem in the absence of treatment, the no-treatment control can be deleted.

Despite the apparent importance of the placebo condition, a variety of issues have emerged regarding the conceptual, methodological, and ethical problems inherent in placebo designs (O'Leary and Borkovec 1978). Conceptually, the definition of placebo is based on contradiction and ignorance. A placebo effect involves change that is produced by a procedure with no theoretical reason for being change inducing; the nonspecific condition is one that is designed procedurally to have no active ingredients. Yet, control for its effects is required out of recognition (Shapiro and Morris 1978) that nonspecific factors do produce change (routinely superior to no-treatment conditions), albeit through poorly understood mechanisms. Perhaps the best description of the placebo condition, then, is that it involves contact with a therapist who engages in methods that the client believes will be helpful, even though the therapist (or investigator) believes that the method will be of only limited effectiveness relative to the therapy condition to which it is compared. Whatever active ingredients it contains are common across many forms of psychosocial therapy, but the mechanisms of its action are a mystery (and, incidentally, a potentially crucial area of investigation in its own right).

In the abstract, this is not a problem for design: The placebo and therapy groups potentially contain the same degree of these active, nonspecific factors, so that differential outcome can be attributed to what is specific and unique to the therapy procedure. The pragmatic design problem is, however, at least fourfold. First, it is not possible to design a placebo procedure that is inert by all theories. Second, unless the placebo condition generates as much credibility and expectation of improvement in the client as does the therapy condition, it is not serving its control purpose. Therefore, the investigator is faced with the impossible task of creating a condition that is inert by all therapy theories and the difficult task of generating its procedures in such a way that clients will find it believable and have faith and hope equivalent to that induced by actual therapy methods. Third, placebo conditions that are created often bear little

resemblance to the therapy condition; thus, the two procedures vary on many dimensions, making it difficult to rule out numerous rival hypotheses in explanation of observed differences.

It is useful to remember that the placebo methodology was adopted by psychotherapy researchers from medication designs. Drug placebos can be matched to the active medication very precisely on such variables as size, color, taste, and side effects; they differ only on the specific active chemical ingredient, and double-blind methods insure protection from differential expectancy effects in both client and drug administrator. The scientific ideal of such a degree of match with psychotherapy methods is extremely difficult to achieve. Finally, although the problem of therapist bias is not fully resolved by control conditions to be discussed later, the potential confounding impact of therapist expectation is particularly acute in the case of placebo conditions. It is difficult for therapists to give to a client a procedure believed to be inert. Therapist expectation, comfort, and enthusiasm are quite likely to vary considerably from those associated with active forms of treatment and may therefore represent a confounding to the interpretation of differential outcome.

It is possible to consider ways of dealing with some of these problems so that placebo conditions might still be validly used. First, placebos should not contain procedures that major theories about the diagnostic problem would suggest are active. Second, placebos may be best developed if they involve procedures that the theory of the therapy condition being evaluated in the study suggests are the usual processes of maintenance of the problem. Such a theoretically relevant placebo condition is nearly ideal from a conceptual standpoint, since it is, by the theory of the therapy to which it will be compared, definitely inert. Third, to the degree that placebo conditions are similar in procedural details to the active therapy being investigated but without the crucially active combination of procedures, the likelihood that equivalent credibility and expectancy will be generated in the client is increased. Fourth, the credibility and expectancy of the placebo method can be assessed on a pilot sample of the relevant diagnostic group (as well as on the clients in the actual study at an early point in therapy) to evaluate whether it generates equivalent levels of these psychological processes relative to the therapy condition, and modifications in its procedures can be made until it does. At present, there is no consensual method of dealing with the problem of therapist bias when placebo conditions are used.

While some of the above methodological problems can be handled, an ethical problem exists when a procedure predicted to be noneffective is administered to people seeking help; its presentation to a client places the investigator in at least a marginally deceptive position. The justification for doing so rides on the equally important ethical requirement that therapy procedures offered to the public be validly evaluated and on the empirical

fact that placebo procedures do indeed produce significant change. For these reasons, placebo control continues to be used. Special safeguards need to be implemented, however. As with no-treatment participants, clients in placebo groups need to be monitored for deterioration during the trial and removed from the study if this occurs. Second, although this will be the case for any client failing to benefit from participation in the project (as will be discussed later), it is particularly important to provide placebo clients with an opportunity to receive an active treatment at the end of the therapy trial, even though this may mean the loss of followup data. Finally, the nature and course of the disorder and the duration of the treatment trial need to be considered before the decision is made to use a placebo condition. More serious disorders that require active intervention, disorders known to be unresponsive to placebo treatments, or longer durations of treatment trial would argue against the use of such conditions.

Although it is customary in medication trials to inform prospective clients that they may be assigned to a placebo condition, this informed consent procedure is not routinely employed when psychosocial placebo groups are used. The effectiveness of the latter rides so much on clients' expectations that they are receiving useful therapy that informing them of the possibility of inert treatment is seen as seriously deleterious to the possible effects of all conditions employed in the study.

Other problems exist about which we know very little. Clients participating in a condition that is minimally effective may become less optimistic about their disturbances, have less faith in the mental health profession, or fail to seek more useful resources that may alleviate their problems. The continued use of placebo conditions thus places a strong responsibility on the investigator to safeguard the welfare of his/her clients and on researchers in general to develop alternative methods of ruling out nonspecific effects.

Best Available and Treatment-As-Usual Comparison Designs

Because of such ethical problems with the placebo condition, some investigations have employed conditions that represent standard approaches to the disorder against which to compare the therapy condition. "Best Available" conditions involve therapeutic methods that are frequently used in mental health settings for a particular diagnostic problem, even though there may be no empirical evidence that they are active therapies. The assumptions are that they are at least minimally effective, given their common use, and that they should also be able to serve as an adequate control for nonspecific effects. The former thus allows a conservative estimation of the potency of the therapy condition being evaluated, since it is a potentially active intervention, while the latter is the case as long as credibility and expectancy are found to be equivalent for the best available condition and the therapy method under investigation. The disadvantage of this control group is that it is likely to vary in many ways from the therapy condition,

BEST COPY AVAILABLE

thus allowing conclusions about efficacy but adding little to basic knowledge about the technique's active ingredients. "Treatment-as-usual" has the same advantages and disadvantages; it typically stands for the various treatment methods customarily employed at an inpatient facility. Thus, an investigation may compare whatever package of therapy procedures is routinely offered to a diagnostic group at the hospital to the new therapy technique under evaluation. The potential for a special methodological problem with this approach exists, however: If clients are in close proximity and become aware of differential treatment, resulting confounding factors may influence the validity of the data.

Component Control or Dismantling Designs

Continuing to pursue the strong inference approach, if a therapy has been demonstrated to be superior to a nonspecific control condition, we can then take what is left over, the specific therapy procedures themselves, and identify some of their individual procedural elements. The Component Control or Dismantling Design randomly assigns clients to conditions each of which contains a distinct subset of therapy elements derived from the package as a whole and contrasts these with each other and ordinarily with the total package. To give an historical example, component control investigations of systematic desensitization therapy have often employed the following conditions: relaxation training alone, repeated imaginal exposure to a hierarchy of feared scenes alone, and the combination of these two elements (systematic desensitization in its traditional form). This game is very subtle, since there is a control condition not yet mentioned, often ignored in past research of this type on the technique, and yet extremely important from the point of view of the original theory of desensitization. The theory (Wolpe 1958) postulated the importance of counterconditioning in fear reduction, procedurally implemented through the contiguity of deep relaxation and repeated exposure. Thus, a potentially crucial control condition, if one is interested in this theoretical question, would involve a group that receives both relaxation training and repeated scene exposures, but noncontiguously. Thus, clients in such a condition might undergo imagery exposure during the first half of the session and relaxation training in the second half. Assume that the traditional technique is superior to the component control. Imaginal exposure and relaxation training were common elements to both methods, so we can rule out the mere presence of both elements as an explanation of the change. What is left as the crucial, specific ingredient to explain the change is the contiguous presentation of imaginal exposure during a deeply relaxed state. Selection of components for a design can thus be based on a consideration of the individual elements of procedure that can be identified, on the theoretical foundation underlying the therapy technique, or on both.

One crucial feature of the component control design is that more factors are ordinarily common among the various comparison conditions. In addition to

representing equally the potential impact of history, maturation, and so on and the impact of nonspecific factors, a procedural component is held constant between the total package and the control condition containing only that particular element. Such a design approximates more closely the experimental ideal of holding everything but one element constant. Differences between the conditions allow the ruling out of the main effect of that particular component as the sole cause of change. Whatever is left over in the total package (the uncommon elements between the conditions) remains as the hypothesized site of active ingredient and the unrejected aspect of the procedure to be targeted for future pursuit. Thus, progress can continue toward determining increasingly specific cause-and-effect conclusions. At the theoretical level, such outcomes tell what elements of procedure are most actively involved in the change process and thus reduce the possible, likely theoretical interpretations of how change comes about, directing attention to these unrejected components for further pursuit. At the applied level, determination of elements that do not contribute to outcome allows therapists to dispense with their use in therapy.

Several additional benefits are inherent in the component control condition. Therapists will usually have greater confidence in, and less hesitancy to administer, a component condition than a pure nonspecific condition. They will also be equivalently trained and have equal experience in the elements relative to the combination of elements in the total package. Finally, the ethical issues are not quite as severe. Any one of the elements may be responsible for change in the total package, so offering an element is not as deceptive as in the case of placebo conditions. While there is much to say in favor of the component control condition, its major drawback is that not all therapy methods may be susceptible to a component analysis. Cognitive and behavioral therapies can be broken apart into elements or combinations of elements fairly readily, but other forms of therapy may not be as easily open to this approach.

Additive or Constructive Design

An allied design approach is called the Additive or Constructive Design, since an investigator may add two distinct therapy procedures together into a total package and contrast the package with each of the individual procedures. The goal is ordinarily to develop an even more potent therapy based on empirical or theoretical information that suggests that each therapy has reason to be partially effective, so that their combination may be superior to either procedure by itself. In terms of design advantages and possible conclusions stemming from the design, the component control and additive approaches are similar. It is partly the direction of reasoning of the investigator and the history of literature associated with the techniques and the diagnostic problem that determine which design strategy seems to be taking place.

Parametric Comparison Conditions

An extremely sophisticated experimental approach is represented in the parametric design. Here, a therapy procedure, or one of its components previously identified to be crucial, is varied across comparison conditions along some dimension considered to be theoretically important. Comparison conditions are created that represent points along the chosen dimension. Some examples are: duration of exposure to phobic stimuli, duration of tensing muscles in progressive relaxation training, degree of warmth and unconditional positive regard, depth of emotional processing during experiential therapy approaches, and frequency of interpretation. As usual, both applied and theoretical conclusions emanate from the design. For applied purposes, such research can identify the optimal point along the dimension to be used in therapy to achieve maximally certain effects. From a theoretical point of view, such research is defining a function along a continuum and therefore tells us something basic about the underlying principles of human behavior exemplified within this research domain. As in the Component Control Design, each parametric condition provides a potential nonspecific control for the others and allows for the potentially equivalent operation of history, maturation, and so on. The parametric design comes closest to the scientific ideal: All is held constant except for one thing, and that thing is varied systematically along a single dimension.

Comparative Design

The final, general design to consider involves the comparison of two unique approaches to therapy. This Comparative Design sets up something of a horse race between two established therapies, each of which is considered from differing theoretical orientations to be effective for the diagnostic problem. Significant differences between the conditions allow a possible determination of the most effective form of intervention—a potentially significant outcome for applied considerations. At the basic design level, differential outcome also provides a potential opportunity to rule out nonspecific factors as the sole contributor to change, since the inferior method presumably includes those factors to an equal degree (though, again, this must be verified by credibility/expectancy checks). There is a very serious limitation, however, to the likely validity of conclusions about comparative efficacy and thus to the overall usefulness of the Comparative Design. Valid conclusions here would require that each form of therapy be administered with the same degree of quality and expertise. This knotty issue will be discussed at greater length later. Finally, there is very little offered by the design that bears on theoretical conclusions. Two therapies are likely to vary from each other on numerous dimensions; thus, we are far afield from holding everything but one element constant. Other than nonspecific factors, there is little procedurally common between them that can be ruled out, and many uncommon features between them may be the source of any observed significant difference. Interpretations having to do

with why one technique is superior to another are left without experimental evidence.

ADDITIONAL METHODOLOGICAL CONSIDERATIONS

Several additional methodological considerations are important to consider, irrespective of type of comparison conditions employed. First, a therapy procedure that is to be evaluated must in its operational definition be faithful to the technique it represents, and its conduct in a particular study must be replicable. The main vehicle for judging how adequately the study meets these two requirements is a detailed manual of procedure that can be made available to readers. This manual will be used by therapists in the study to insure adherence to the protocol as it is intended. The manual needs to be of sufficient detail so that readers are in a position to know whether the spirit of the technique is being fulfilled and whether all of its methods are included and will be properly implemented. The same detail is required for any other control condition used in the design for all of the same reasons.

The manual will also serve as a resource for developing a system of integrity checking required to insure that during the trial, each of the different conditions is administered in an accurate and nonoverlapping fashion by the therapists. Ordinarily, a checklist of specific therapist interventions is used—one that indicates clearly those methods that are allowed by the protocol of a particular condition and those that are not. Such integrity checks are primarily concerned with insuring that interventions distinctive of one condition do not occur in any other. These integrity checks are usually applied to a random sample of at least 20 percent of all sessions. They need to be conducted by independent raters, who score the session against the checklist from tape recordings of the sessions. The investigator uses the output of these integrity checks as well as his/her own listening to session tapes to monitor adherence to the protocol by each therapist in each condition and to take steps to correct any deviation from allowed procedure. Too many deviations would invalidate the data of those particular clients, although a consensual, operational definition of "too much" has not been established.

A related notion is that of therapy quality, or how much expert administration of the therapy is taking place in the sessions. This issue is most strictly relevant to the conduct of actual therapy approaches and their components and has less clear relevance to nonspecific conditions. Judgments of quality have their greatest importance in comparative designs and some additive designs. There is currently no consensual resolution on the judgment of quality. It is obviously important to a fair comparison between two therapies that each therapy be administered with matched and hopefully high quality; otherwise, a confounding exists. But how to measure quality of any given therapy approach, much less how to define operationally

comparative quality between two highly different therapy traditions, remains an unanswered question. There has been recent movement toward scale development for quality assessment for some therapies, and many investigators employ expert consultants to listen to randomly selected session tapes and make judgments of quality with or without such formal scales. But at the present time, there is no standardized procedure beyond finding the best therapists one can and training them in a well-detailed protocol manual that faithfully represents the therapy orientation and interventions being investigated.

Prior to launching a large-scale therapy outcome investigation, it is incumbent on the investigator to conduct pilot applications of each condition. Without a pilot trial, it remains uncertain whether the therapy as "manualized" can be administered properly or whether clients will find it acceptable as a form of treatment for their problems.

A valid outcome investigation must employ more than one therapist. A single therapist conducting all therapy procedures confounds treatment effects with therapist characteristic and bias effects and with their interaction with the different comparison or control conditions. It also limits the generality of any findings to the known and unknown characteristics of that particular therapist. When it is possible to have therapists equally competent in the administration of each condition in the study, then it is ideal to have multiple therapists conducting therapy with an equal number of clients in each condition. External validity is increased, and the likelihood that therapy condition effects are confounded by therapist characteristics is decreased. With component control designs, equivalent therapist competence is easily accomplished. With comparative designs and some additive designs that make use of therapies from differing theoretical traditions, it is not so easy. In general, faithful representation of a technique with high quality cannot be insured when it is conducted by therapists previously untrained or relatively inexperienced in that technique. A confounding results if the therapists' prior training and experience have primarily been with only one of the techniques in the study and they have little or no knowledge of another comparison condition, beyond the training provided for the sake of the study. Differential outcome could be due to one of the conditions truly being superior or, alternately, to better administration of, or therapist belief in or enthusiasm for, the condition that produces the better outcome. This problem cannot always be circumvented. When contrasting, say, short-term psychodynamic therapy with experiential therapy, it is nearly essential to have one set of therapists well trained in the former technique provide only that therapy while therapists trained in the latter tradition administer only that condition. Faithful and high-quality representation of each therapy is so essential that the investigator may be willing to allow the confounding, but interpretation of the result will be less clear. The differential outcome could be due to the therapy, to this set of therapists, or to their combination. Whether to proceed with such an

investigation is partly a function of the state of the literature at the time and whether there are other design approaches that will be pursued later that might ultimately decide about the relevance of the therapist confounding.

Given this variety of issues having to do with the therapists in a project study, it is important for the investigator to specify carefully the therapists' characteristics, their training background, and their experience level in clinical work in general and in the techniques of the study in particular. Moreover, some detail should be provided that indicates how they were trained in the specific protocols prior to the onset of the project and how they are supervised during the conduct of the trial.

All conditions included in a study must be matched as closely as possible to each other to reduce the number of ways in which the various conditions might differ from each other and thus the number of rival hypotheses that could explain differential outcome. Most importantly, this includes the amount of contact with a therapist. Consequently, all clients, irrespective of condition, are seen in therapy for the same number of sessions and the same length of time in each session. Any other procedural aspect of the primary therapy condition under investigation that can reasonably be included in control conditions should be. For example, if homework assignments of a certain duration are given in the therapy procedure, then some type of homework of equal duration should be part of the control-condition protocol.

A potential problem regarding the matching of therapy time arises when using component control or additive designs. If a control condition involves a component of a therapy package or one of the several interventions in an additive design, then the investigator is faced with a decision that is most clearly seen in an example. Assume that a two-component therapy package to be evaluated requires 12 60-minute sessions, with 30 minutes devoted to component A and 30 minutes devoted to component B. The study will contrast the entire package (A+B) with two control conditions, one representing each of the components. For each of these control conditions, does the investigator provide 12 30-minute sessions, thus confounding outcome comparisons by differential amount of contact with a therapist, or are full 60-minute sessions given of the component treatment, thus providing twice the amount of therapy with that element that clients in the total package receive? There is no absolute consensus on resolving this issue. The currently held ideal is that nonspecific filler procedures should be added to component conditions so that both the total amount of therapist contact and the amount of therapeutic attention devoted to any element is held constant over all conditions. Thus, neither confounding is present in the design. Of course, all of the caveats about nonspecific control discussed earlier apply here, only with ordinarily much less methodological and ethical concern. Other solutions simply allow one or the other of the two confoundings (total time or total amount of treatment by a particular

component), in which case discussion of the results must clearly indicate that these rival interpretations remain. The best way to use these latter solutions is to conduct the study twice, once confounding total time and, upon replication, confounding total amount of treatment. If similar differential outcomes emerge, both confounding factors become less likely rival explanations. This is, however, an obviously costly solution.

It is very important to insure that assessment and data management staff remain unaware of the condition status of clients throughout the trial and followup periods. Otherwise, subtle influences based on that knowledge could influence the results. Often, clients are requested at the beginning of an assessment session not to discuss any aspect of their therapy participation with the assessors in order to decrease the likelihood of clients disclosing condition-identifying information.

It is important to do a power analysis before proceeding with an outcome investigation (Cohen and Cohen 1975). The power analyses will provide the number of subjects required in each condition in order to detect a difference between conditions, given an estimated effect size and the probability level that the investigator wishes to establish for detecting an effect that large. The effect sizes are not always easy to determine; prior research using a specific therapy with the relevant client sample is not always available. Such data should be used if they do exist, but otherwise one must guess whether differences are likely to be "small," "medium," or "large." The more effective a control condition is in holding things constant (e.g., a nonspecific condition that has been shown to be markedly effective with this particular client group, a component control condition that is very similar to the therapy condition, or parametric conditions), the smaller is the likely true difference between it and the therapy condition (if indeed a true difference exists), and the larger the number of subjects required to show a significant difference.

Random assignment to conditions is, of course, essential. Otherwise, the internal validity of the experiment is threatened by differential selection, differential attrition, and interactions between selection factors and history, maturation, and so on. Because certain variables are likely, or are known, to relate to outcome with certain client groups (e.g., age, gender, presence or absence of medication, and pretherapy severity of disorder), however, investigators commonly block on such variables and randomly assign within blocks to insure pretest equivalence among the conditions. It is also essential to analyze all pretherapy demographic and dependent variables to demonstrate that the conditions were equivalent prior to therapy. The presence of pretherapy differences introduces a potential confounding. Although covariance analysis can statistically correct outcome scores (using involved variables as the covariates), results remain suspect, since the groups were psychologically different from the start.

Because therapy-outcome investigations tend to take place over a long period of time, with clients being recruited in waves, the investigator must be sure that clients are being assigned proportionately to each of the conditions at any given time. If an unbalanced number of clients are seen in one particular condition, this would introduce several sources of potential confounding, such as sessional effects and amount of experience among therapists and assessors over the total duration of the trial.

Attrition, if in large numbers or differential between conditions, will invalidate the results of a study, since it potentially destroys initial equivalence among conditions. It is to be hoped that random assignment will balance the conditions with clients who are likely to drop out, but certain conditions (e.g., placebo, no-treatment) may by their nature encourage a larger dropout rate relative to actual therapies. Assuming low or not markedly differential attrition (tested by chi-square) among the conditions, there are two further steps that should be taken. First, efforts should be made to obtain postassessments and reasons for dropping out from clients who are terminating near the point of their last contact with the project. Outcome analyses are then conducted twice, once without dropouts and once with dropouts, using their last obtained data for any assessment points. Second, dropouts need to be compared to completing clients on pretherapy demographic and dependent variables. To the extent that these analyses remain relatively unchanged irrespective of who is included and show no differences between remainers and dropouts at pretherapy, the potential biasing effect of attrition can be assumed to be minimal.

It is critical that the investigator establish a priori operational definitions of treatment responders, failures, and followup relapsers so that proper continuing care can be provided to clients who have not benefited from their participation in the project. This is particularly true for clients assigned to any form of treatment considered to be minimally effective but may be relevant even to some clients given that therapy felt to be the most effective. These operational definitions are usually in the form of failure to achieve 20 percent or more improvement (or a return below this level at followup) on two of three, or three of four, major outcome measures. Further treatment for such failures or relapsers, with the best performing condition in the trial or with some other form of individualized clinical therapy considered to be beneficial for clients with this diagnostic problem, is then provided. This may result in clients lost to subsequent followup, but again ethical requirements take precedence.

Confoundings to conclusions can enter over the course of a therapy trial because clients may seek alternative modes of treatment, sometimes during the trial itself but most often during followup periods. Clients participating in outcome trials are not ordinarily entered if they are being seen in therapy elsewhere at the same time, and they are entered with the understanding that no other therapy will be sought while they are being treated in the

trial. Investigators, however, cannot ethically place such a restriction on clients during the often lengthy followup periods. It is important, therefore, to assess at followup whether clients have received additional treatment since their participation. Differential amounts of such therapy over the different conditions would raise a rival hypothesis regarding followup results.

CONCLUSIONS

In the beginning of this paper, two goals were mentioned for therapy outcome research: (1) demonstration of efficacy of a therapy technique and (2) identification of the mechanisms of the technique by which it facilitates change. Although comparison to a nonspecific condition would provide evidence that an intervention has a specific, active ingredient and thus could be considered to have useful efficacy, it is important to realize the rather limited nature of such a demonstration. Research cannot usefully stop at this stage. The ultimate goal of therapy development and evaluation research is the identification of interventions that are 100 percent effective, a mark not even nearly reached for the vast majority of disorders. In the long run in technique development, it is essential to build basic knowledge about a technique's mode of action in order to learn more about the nature of the disorder and the nature of the principles operative in the (even partial) effectiveness of the therapy. Thus, the question "Is this therapy effective with this disorder?" is only a preliminary step in the direction of truly useful research. Moreover, from the above discussion, it is clear that conclusions from therapy/placebo comparisons will remain necessarily vague due to the variety of problems associated with nonspecific control conditions, relatively uncontrollable therapist biases, and so on. Component control and parametric designs suffer from fewer methodological problems and approximate more closely the matching of conditions on numerous dimensions, allowing the investigator greater specificity in cause-and-effect conclusions. These designs contribute significantly more in basic knowledge because of this specificity, a specificity that narrows the possible theoretical interpretations of possible mechanism and stimulates the generation of novel interpretations within the restraints of the results yielded by these designs. But beyond these comparison-condition considerations, progress in basic knowledge from therapy outcome studies can be considerably accelerated to the degree that each study incorporates certain methodological and measurement options that bear on issues of mechanism. These generally fall into two major domains: (1) factorial designs that incorporate personality or subject characteristics predicted by theories of mechanism of change to relate to responsiveness to different types of therapy or their components and (2) acquisition of auxiliary pretest and/or posttest measures and/or during-therapy process measures predicted by such theories to correlate with outcome. Knowledge will thus accumulate more rapidly to the extent that the strong inference approach, leading to increasingly specific cause-and-effect conclusions, is applied not only to therapy procedures but

also to client variables, therapist variables, outcome variables, and process variables. In each case, one takes whatever was unrejected and breaks it into separate elements for evaluation. It is this process that will most efficiently lead to answering the ultimate therapy question "What kinds of therapy delivered by what kinds of therapists are effective in the short term and the long term, reflected by what breadth of changes with what kinds of people suffering from what kinds of disorders, and how do those changes come about?"

REFERENCES

Campbell, D.T., and Stanley, J.C. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand-McNally, 1963.

Cohen, J., and Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale, NJ: Erlbaum, 1975.

O'Leary, K.D., and Borkovec, T.D. Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *Am Psychol* 33:821-830, 1978.

Paul, G.L. Behavior modification research: Design and tactics. In: Franks, C.M., ed. *Behavior Therapy: Appraisal and Status.* New York: McGraw-Hill, 1969.

Platt, J.R. Strong inference. *Science* 146:347-353, 1964.

Shapiro, A.K., and Morris, L.A. Placebo effects in medical and psychological therapies. In: Garfield, S., and Bergin, A.E., eds. *Handbook of Psychotherapy and Behavior Change.* New York: Wiley, 1978.

Wolpe, J. *Psychotherapy by Reciprocal Inhibition.* Stanford: Stanford University Press, 1958.

ACKNOWLEDGMENT

Preparation of this report was partly supported by the National Institute of Mental Health Research grant MH-39172.

AUTHOR

T.D. Borkovec, Ph.D.
Professor of Psychology
The Pennsylvania State University
University Park, PA 16802

Attrition in Substance Abuse Comparative Treatment Research: The Illusion of Randomization

Kenneth I. Howard, W. Miles Cox, and Stephen M. Saunders

INTRODUCTION

There have been two broad approaches to establishing the relationship between the domain of empirical data (facts) and the domain of theoretical ideas (explanations). One approach—the hypothetico-deductive method applied to formal theory—emphasizes the primacy of theory and adduces evidence to either support or refute theoretical hypotheses. It is exemplified by Hull's (1943) work in the area of learning theory. This method has been described as the **confirmatory approach**, even though a theory cannot be confirmed and the best that we can do is to reject the null hypothesis. The other approach emphasizes the primacy of data and seeks to provide explanations for observed patterns of phenomena; the observation of patterns leads to concatenated theory (Kaplan 1964). This approach has been described as the **exploratory approach** and is exemplified by Skinner's (1938) functional analysis of behavior.

Each of these approaches has advantages and disadvantages, but there is a natural tension between them. The confirmatory approach emphasizes internal validity (Campbell and Stanley 1963) at the expense of the generalizability of findings. It invokes the assumption of *ceteris paribus*, which is almost always counter to fact in clinical research, in the search for verification of theoretical hypotheses. This approach entails a sufficient condition methodology that seeks to demonstrate how things **could** happen.

The exploratory approach emphasizes generalizability (external validity) at the expense of eliminating alternative plausible explanations for the results (internal validity). It admits the influence of a large number of potentially confounding variables and seeks to eliminate that influence through the use

of (imperfect) statistical analyses. It seeks to establish how things do happen at the expense of fully specifying a causal statement.

Comparative treatment research has evolved in the context of confirmatory methodology, and the following discussion will focus mainly on this approach.

CONFIRMATORY METHODOLOGY

Following Fisher's (1925) introduction of the analysis of variance and its associated methodology, Underwood (1949) and others further explicated the value of true experiments for research in psychology. Underwood's influential treatise helped to shape the "official" methodology for empirical inquiry in the social and behavioral sciences. The sine qua non of this methodology is the random assignment of objects to conditions (levels of the independent variable). Such random assignment is intended to ensure that any observed differences in the dependent variable, across levels of the independent variable, can be attributed solely to the influence of the independent variable and not to the influence of preexisting differences among the objects. Hence, a causal relationship between an independent and a dependent variable can be inferred. Statistical methods were developed, moreover, to ensure that any observed difference between conditions was not due to chance. These statistical methods provide standards for rejecting the null hypothesis of no real difference. With random assignment, it is thought, chance is the only rival explanation for observed differences. In the absence of random assignment, however, preexisting differences between groups, subject variables, and other "uncontrolled" variables become rival explanations.

As noted above, confirmatory methodology has been adopted as the model for comparative-treatment research. The independent variable in this research is usually two or more treatment conditions for a class of patients randomly assigned to each of those conditions, and the dependent variable is usually improvement of the relevant clinical syndrome. Two problems arise when this methodology is applied to patient populations. First, because of the multitude of potentially relevant variables, the sample size for each treatment is never sufficient to ensure that random assignment will equate groups on all relevant subject variables, i.e., will satisfy a *ceteris paribus* clause. Thus, even when groups are "equated" and significant outcome differences between them are obtained, we cannot be sure that these differences are solely a result of the causal influence of the selected independent variables. This limitation makes replications absolutely essential. (For an instructive example, see Lovibond and Caddy 1970; Caddy and Lovibond 1976; Lovibond 1975.) Second, in conducting treatment research with patients, it is virtually impossible to avoid missing data (attrition) because patients routinely fail to provide complete information at all data points and routinely fail to complete treatment

regimens. Missing data, therefore, always compromise the equating of groups across treatments.

THE PROBLEM OF MISSING DATA IN COMPARATIVE-TREATMENT RESEARCH

Attrition entails the nonavailability of data for certain patients at certain assessment points in a study. Traditionally, attrition has been considered a problem only after a sample of patients has been selected for a project, but an examination of the basic issue of data loss suggests that attrition may also be problematic at other points in the execution of a study. In fact, loss of relevant patient data can take place (1) before patients enter the study, (2) at intake, (3) during the initial screening, (4) after assignment to a treatment, (5) at various assessment points during treatment, (6) at the completion of active treatment, and (7) at assessment points during the followup period. Even a casual survey of the substance abuse treatment research literature shows that these various sources of losses are very common indeed. For example, in a recent comprehensive review of 68 studies of the treatment of alcohol problems, Riley and colleagues (1987) found that an average of 24 percent of patients who completed treatment were lost to followup.

Serious damage can be done when a soundly designed study is subjected to the exigencies of subject recruitment and data collection. The fourth point noted above, assignment to a treatment (the independent variable), forms a natural division for thinking about loss of data. That is, the first three points define the **preinclusion** phase of a study in which potential subjects and their data either do or do not become available for the study. The last three points define the **postinclusion** phase, in which selected subjects either do or do not remain available to the study, providing complete dependent-variable data. Thus, we delineate our discussion according to **preinclusion** attrition, **independent-variable** attrition, and **postinclusion** or **dependent-variable** attrition.

PREINCLUSION ATTRITION

The inclusion of subjects in a comparative-treatment research project depends upon natural constraints on the pool of available subjects and on decisions made by both the investigator and the prospective patients. With regard to constraints on the availability of subjects, we already know that only a minority of persons with substance-abuse problems present themselves for treatment and those who do so are not representative of the entire population of substance abusers.

In addition, investigators themselves set certain selection (inclusion and exclusion) criteria, e.g., age, sex, chronicity, diagnosis, or suitability, either for reasons of relevance to the problem under study, for ethical reasons, or

for the sake of convenience. Moreover, prospective patients are not passive respondents to invitations to participate in a study. They make their own assessments of the probable benefits and costs they may sustain and agree or decline to participate for their own reasons. For example, they may not think that the kind of treatment offered is relevant to their condition or consistent with their values. Further, their motivation to volunteer for research may be related to their motivation to recover (Cox and Klinger 1988) and hence interacts with the treatment condition to which they are assigned.

These decisions by prospective research subjects are not necessarily made at one point in time. When patients are first made aware of a study, they have only a limited knowledge of what will be expected of them if they decide to participate. If their initial impressions are favorable and they proceed to the intake procedure, the requirements of the study will be explained to them in greater detail. At that point, some patients will not accept the terms of the research design or will refuse to sign an "informed consent." Others will simply fail to keep their appointments. The demands of a study become more apparent to patients during the initial screening process, when they are likely to be interviewed and/or given a time-consuming battery of tests. Some patients will drop out during this screening procedure, whereas others will provide only partial data.

For example, in the Veterans Administration cooperative study of disulfiram (Antabuse) for the treatment of alcoholism (Fuller et al. 1986), the authors state that 6,629 patients were screened for inclusion in the study. Of these, 5,011 (76 percent) were ineligible because they lived alone, had a contraindicating medical condition, lived more than 50 miles from the hospital, were older than 59, abused drugs, exhibited destructive behavior, had a history of being uncooperative, or had been abstinent for more than a month. Of the original 6,629 patients, 1,618 (24 percent) met inclusion criteria for the study. Of these, 1,006 refused to participate. Thus, of the original sample, 612 (9 percent) were randomly assigned to the three treatment conditions. (Of course, not all of **these** patients provided complete data during the actual study.) There is, in other words, a natural filtering effect through which some members of the population of interest, e.g., alcohol abusers, who are deemed by the researcher to be appropriate to include in a study, e.g., present to an alcoholism treatment unit, become unavailable, either through their own choice or through intervening circumstances. Other members of that population will have a disproportionately greater chance of being included in the investigator's treatment groups. The actual study sample may respond differently to one or the other of the treatments that are being evaluated than would a more representative sample of the target population.

Preinclusion attrition has a subtle influence on research, and one that is not easily recognized by investigators, because the potential subjects have not

yet contributed to the data pool. Although preinclusion attrition is not often reported because it is rarely recognized, it can introduce bias into a study by making the study sample in some ways unrepresentative of the population about which the investigator wants to learn. At the very least, an investigator should take great care in describing the patient population for which any findings may be generalized. Every research report should include a statement describing the parent population, e.g., the incidence and characteristics of cocaine abusers in Chicago, the fraction of that population who come to the attention of the project, the fraction who meet inclusion criteria, and the fractions who remain after each "hurdle," e.g., signing an informed consent. It is not only incumbent on the investigator to describe the parent population and the sample drawn from it, but it is also the investigator's responsibility to explain the impact that this natural filtering (preinclusion attrition) may have had on the results of the study.

INDEPENDENT VARIABLE ATTRITION

In comparative-treatment research, the independent variable is the set of treatment conditions to which patients are assigned. Some patients will not actually receive the assigned treatment, i.e., some patients will not have any exposure to the treatment, whereas others will not complete the specified treatment regimen, i.e., some patients will not be adequately exposed to the assigned condition. In order to ascertain the extent to which application of the independent variable had the intended impact, "manipulation" checks and treatment-integrity checks are essential, although they are not routinely carried out in substance-abuse research. In a study in which the investigators did check the effectiveness of the independent variable (Sanchez-Craig et al. 1987), they unexpectedly found it to be quite weak. In this study, alcoholic patients in a halfway house were taught cognitive/behavioral skills to deal with negative emotions that promoted their tendency to drink. Even though they had mastered the skills during the treatment program, they tended to forget them within 1 month following the completion of treatment.

The most common way to deal with attrition in a study is to ignore it. It is not unusual for an investigator to report the characteristics of the sample that was randomly assigned to the various treatments and not to mention attrition from the various cells of the design. The reader is left to try to make sense of the shifting N's used for each data analysis. A variation on this practice is to report the pattern of attrition (the number of patients surviving each filter or hurdle) and then to proceed with the data analyses with no attempt to deal with the impact of this attrition. In either case, it is rare indeed that an investigator will report the characteristics of the patients on whom the data analyses were actually based.

Another common appr .n is to replace defecting subjects. This maintains the sensitivity and pc .er of statistical contrasts. In terms of

generalizability, this is no different from basing the analyses on the panel of patients who actually completed treatment in the first place, since it is merely an enlarged group of such patients. All that the replacement tactic provides is a larger sample size of treatment completers and the illusion of random assignment provided by the appearance of equal cell sizes.

Still another approach for dealing with independent variable attrition is to compare the patients who receive the treatment with those who do not with regard to other variables on which measures are available, e.g., age, gender, or education. If there are no statistically significant differences between these two groups on any of these measures, then the investigator will conclude that the design was not seriously compromised by attrition. There are several problems with this approach: (1) The sample size for comparing completers and "attritors" is usually not large enough to provide adequate power for detecting differences between them (particularly if this comparison is made within treatment conditions); (2) The variables used in the comparison are usually not directly relevant to the clinical condition under study; and (3) Most importantly, the failure to detect statistically significant differences does not imply group equivalence and certainly does not warrant acceptance of the null hypothesis. This approach, when it fails to show differences, ultimately assumes that attrition was random on any causally relevant variables across the treatment groups. It is then used to support the validity of an analysis of the data provided by the treatment completers alone.

Yet another common approach is end-point analysis. In this case, **assignment** to a treatment condition becomes the treatment whether or not a patient ever was exposed to the treatment. This approach maintains the integrity of the random assignment at the expense of comparing treatment outcome—it puts method ahead of substance. It does, however, have relevance for service-delivery research, where the attractiveness of a treatment (Moos and Finney 1983)—and the cost of implementing it—is also a measure of its effectiveness. From this point of view, a treatment that is very effective but is not acceptable to patients might be seen as worse than a treatment that is moderately effective but attractive.

One problem with end-point analysis is that some attritors will be considered successes, and this may artificially inflate the apparent effectiveness of a relatively unattractive treatment. A patient can drop out of an assigned treatment condition and seek treatment elsewhere or otherwise achieve recovery—there are many therapeutic events in life. If a condition engenders a high attrition rate and some attritors are nonetheless ultimately successful, end-point analysis might lead the investigator to conclude that the treatment was more effective than warranted. For example, suppose that any attritor has a one-third chance of recovery due to exposure to other therapeutic events. If treatment A has a 50-percent completion rate, employing end-point analysis, it would ultimately be

BEST COPY AVAILABLE

credited with an additional 17-percent recovery rate (one-third of the attritors). If this treatment, when completed, is very successful, e.g., a 90-percent recovery rate for completers, then it would appear to have a 62-percent success rate. By contrast, if treatment B had an 80-percent completion rate and a 70-percent recovery rate, its success rate would appear to be 63 percent. The conclusion from an end-point analysis would be that the two treatments are equally effective. If, on the other hand, the label "success" were restricted to patients who are successful only after completing the assigned treatment (and dropouts were considered treatment failures), treatment A would show a 45-percent success rate and treatment B would show a 56-percent success rate.

In any case, a research design imbalanced by attrition from treatment will yield uncorrectably biased estimates of statistical aggregates, such as means or components of variance. Attrition that is not outcome orthogonal in the various cells of the design at which it occurs cannot safely be ignored. It likewise cannot be sensibly corrected unless we already know how it is causally related to the dependent variable, cell by cell, or unless its effects luckily cancel out each other in the statistical aggregates of interest. Unfortunately, where there is independent-variable attrition, none of these possibilities can be evaluated on the data. That is, we know only that cases are missing from certain cells, but not which cases from which cells.

POSTINCLUSION ATTRITION

Once patients have been formally included in a study sample, various other factors determine whether they remain in it and contribute complete data for the full duration of the data collection. Therapists may come to feel that the conditions to which their patients have been assigned are inappropriate and may exercise their professional responsibility in a way that effectively removes these patients from the study. Patients, of course, may also decide to discontinue for reasons relevant or irrelevant to their disorders. Patients who drop out of treatment might have less severe symptoms or less need or desire for **formal** treatment than do those who remain. Their dropping out might effectively cause the sample that remains to be more severe or more difficult to treat than the sample as originally constructed, thereby biasing results for or against one or more of the treatments being evaluated.

Patients may remain in treatment but effectively drop out of the study by failing to complete some or all of the measures requested of them during the course of treatment, at termination, or during the followup period. Some substance-abuse patients who were compelled by others, e.g., the court, family, or employer, to seek treatment may not fully participate in the regimen.

Patients may drop out of treatment or may be unlocatable during assessment periods. They may continue in treatment beyond the study period or have

protocol-deviating treatment during the followup period, thereby compromising outcome evaluations at termination or followup. With regard to the latter point, extratreatment influences, such as attending self-help groups, are quite common among substance-abuse patients. These influences and how they might interact with a patient's group assignment in a formal study need to be assessed.

For example, in a study of skill training for alcoholics (Chaney et al. 1978), 10 of the 50 randomly assigned patients dropped out during treatment. Results are actually reported for 37 patients at 3 months and 39 patients at 12 months. Similarly, in a study of partial hospitalization (McCrady et al. 1986), only 101 of the 174 randomly assigned patients provided complete data.

Postinclusion attrition is very readily and often painfully apparent to investigators because they can see data on patients in whom much effort has been invested vanishing from view. Moreover, because of investigators' vested interest in their study and the desire to have the superiority of their favored treatment confirmed, they themselves may have a subtle influence on patient attrition (Smith et al. 1980; Rosenthal 1976). Since double-blind designs (considered indispensable to unbiased evaluations of medications) would be literally impossible to execute in controlled tests of psychotherapeutic interventions, these potential influences can never be ruled out.

When the rates of attrition (the attrition fractions) differ among the various treatment groups, the causes of attrition are also likely to be different, which means that the groups are likely to be no longer truly comparable. Even if the attrition fractions are the same across treatment groups, the comparability of the groups has been lost because similar attrition rates do not necessarily imply similar causation. The effect of postinclusion attrition in undermining the comparability of treatment groups thus constitutes a major threat to internal validity. That is, observed differences between groups on the outcome variables may be mistakenly attributed to treatment effects. Again, to prevent postinclusion attrition from further threatening the external validity or generalizability of the findings, investigators must take particular care to redefine the sample in terms of the characteristics of patients who actually completed the research protocol. It must be acknowledged that the sample, due to attrition, now has special characteristics that are undefined. These are the characteristics that led some subjects to continue to provide data and others to discontinue.

In order to absorb the effects of dependent-variable attrition without bias to one's results, these results can be stated as interval estimates that put bounds on the possible effects of the attrition that did occur. What the results would be like if all lost data were as favorable as possible sets the upper bound, and what the results would be like if all lost data were as unfavorable as possible sets the lower bound (Stephan and McCarthy 1958).

In this approach, a sample of the worst observed values and a sample of the best observed values would be substituted for patients with missing data. For example, in the comparison of two treatments, the attritors from the apparently inferior treatment A could all be assigned a sample of the highest dependent-variable (outcome) scores and the attritors from the apparently superior treatment B could all be assigned a sample of the lowest scores. If subsequent analysis still shows treatment B to be superior, then we can conclude that the results were not reversed by attrition.

Another strategy used to repair the effects of dependent-variable attrition is to substitute cell-mate means for each missing data point. This carries the assumption of random attrition to its logical extreme, assuming that the pooled missing data have, on average, exactly the same characteristics as the observed data. This strategy replaces missing data with fictitious numbers and unfairly accentuates any between-treatment differences by reducing within-cell variance.

Another approach is to assess the extent to which some measured patient characteristics (1) are sufficient predictors of both dependent-variable attrition and dependent-variable levels, (2) are sufficient predictors of any treatment-by-outcome attrition interaction effects, or (3) interact with treatments when their balance across comparison groups has been disturbed (Kalton 1983; Madow 1983). These measurements of patient characteristics can then be used as the basis for repairing the possible effects of attrition on the dependent variables. This could be done by either discarding (or otherwise weighting) cases to rematch treatment groups on these characteristics or using the covariance of these characteristics with the dependent variables to adjust the latter for differences among the comparison groups. This strategy, however, assumes a basic similarity between the attrition sample and the actually obtained sample—a form of random attrition.

A similar approach that is being used more frequently (Mackenzie et al. 1987) is to categorize patients according to the difficulty the investigator has in obtaining dependent-variable evaluations. The patients who were the most difficult to find would form a group of "would have been missing." This group is then compared with other groups to determine whether potential attrition is related to subject characteristics. As before, the assumptions are made that the same factors influence attrition and difficulty and that failure to find statistically significant differences between groups means that the groups are the same.

No matter how carefully a study is designed, the unavailability of certain data is inevitable in every comparative treatment project. As noted above, data will be lost at many points during a study—both before patients have met inclusion criteria and after they have been randomly assigned to treatment conditions. Methodologically, data attrition is of paramount importance because it undermines the equivalence presumably established

among comparison groups by undoing the effects of randomization in the selection and assignment of cases. A variety of methods have been put forth in an attempt to compensate for the effects of data attrition, but all of them can be shown to rest on the untenable assumptions that attritors and completers do not bias estimates of the relationships between independent and dependent variables and that they represent equivalent samples of the same patient population. In other words, they rely on the assumption of random attrition.

ATTRITION PREVENTION

There has been some emphasis on attempts to minimize attrition in comparative treatment studies. Sobell (1978) suggested a five-point strategy aimed at minimizing attrition during the followup period. These steps are: (1) Allow ample time and exert persistence during followup; (2) Before the study begins, provide subjects with information about the followup; (3) Use collateral information resources or contacts; (4) Use official records to locate lost subjects; and (5) Maintain continuity by contacting subjects every few weeks (in the interim between the end of treatment and the followup stage). Gilbert and Maxwell (1987) suggested a sixth strategy that entails identifying subjects who are at risk for attrition, thus alerting the investigator to focus particular effort on following these subjects. These "damage control" strategies have two shortcomings. In the first place, they will be only partially successful, i.e., they will only reduce the attrition fractions, and the remaining data loss will still have to be of concern. In the second place, they alter the nature of the treatments by including continuing contact and, insofar as special attention is given to potential attritors, they compromise the original random assignment.

TOWARD A RECONCEPTUALIZATION

Randomization has been transformed from a method that was intended to assist scientific investigation to a dogma by which research is reflexively judged. We would argue for the opposite stance—**randomization should always be explicitly justified.** An investigator must explain why randomization was undertaken, the extent to which it was successfully implemented, and whether this strategy compensated for the accompanying loss of generalizability.

Along a similar line, there are probably no "main effects" in the real world, in any case, and certainly no main effects in clinical research. The characteristics of any therapeutic phenomenon are jointly, and interactively, determined by the particular patients, the particular therapists, the particular interventions actually involved, and the particular organizational, sociocultural, geographic, and historical contexts in which treatment occurs. Each study of a treatment reflects a limited sampling of patients, therapists, therapies, and contexts. If the comparisons made between different

therapies, different types of patients, different types of therapists, or different contexts are to be internally valid and are to be an accurate representation of differences that would result in the populations sampled (externally valid), then one must find ways to cope with the deleterious effects of data attrition from any of these sources.

We have reviewed strategies for "correcting" attrition, and we have suggested that, except for the most conservative strategy (interval estimates based on substituting best and worst values), they all rely on the untenable and untestable assumption of random attrition. In fact, as we have shown, there are no certain remedies for data loss. Those we have reviewed are limited by the assumptions they require concerning randomness of data loss—assumptions that beg the very question one would want to answer. It appears that we must accept the fact of attrition and realize that we cannot know, much less repair, its effects after it has occurred. We cannot depend on estimates of the counter-to-fact conditional—what would have happened if patients who rejected treatment had accepted it. It is never very satisfactory to speculate about events that could never happen or to come to the conclusion that a treatment would have been effective if only all patients had accepted it. That is a very big **if**.

With or without attrition, one has information about only the variable levels for which observations actually exist. There is no reason to expect these to conform sufficiently to any experimental design to give unbiased estimates of main effects or of interactions among variables. It would be unreasonable to claim, however, that the data collected are therefore worthless. Investigators must accept the inevitable incompleteness of actual observations at the relatively few points they occupy in the independent-variable space defined by the various combinations of patient, therapist, therapy, and setting characteristics. Once this is accepted, a perspective on clinical research opens up that reframes the problem of attrition as a problem of bias. This perspective also, serendipitously, integrates traditional exploratory, single-case studies with the strategy of controlled treatment-contrast, multiple-case research. The basic unit of study in this perspective is the individual case, described with as much dependent- and independent-variable information as is required in the present state of development of substance-abuse research.

In substance-abuse research, what one is really seeking are optimal points in the independent-variable space in terms of dependent-variable outcomes. Optimum-seeking (Adby and Dempster 1974) involves seeking the points in the independent-variable space associated with the best outcomes, and it is surely the most sensible approach to substance-abuse research designed to provide clinically relevant findings. It is the knowledge of which points in the independent-variable space are associated with the best average outcomes that is ultimately of clinical importance. Some sample

methodology and a more technical explication is provided by Howard and coworkers (1986).

CONCLUSION

It is time for us to give up the confirmatory approach to comparative treatment research. In clinical research, we are rarely in the position to test formal, theoretical hypotheses. Clinical research has to be judged by its informativeness, and we have argued that so-called "true" experiments tend to yield trivial information. In any case, it seems self-defeating to espouse a methodology that we must always fail to implement properly and to be forced to move apologetically to secondary analyses to make sense of our results. Instead, we recommend the adoption of exploratory methodology and a greater emphasis on the generalizability and the constructive replication of findings.

Where attrition denies one information at certain points, one must remain ignorant until further cases can be gathered. Where systematically comparable single-case studies yield information on certain specific independent-variable points, something really is learned that, however partial, remains the best one knows until further information is obtained at those points. At certain stages, what is missing may be as important as what is known. The detection of significant gaps in coverage of the independent-variable space serves the useful function of attracting further research attention, especially where optima are suggested to exist. The more extensively the independent-variable space has been covered with cases, the more one can supplement replication at a point by interpolation to that point from surrounding points. Optimum-seeking depends on interpolation and extrapolation from the pattern of observed outcomes across the various points, and precision of estimation at any given point depends on replication. The realistic progress of scientific knowledge requires both optima and precision, and while attrition detracts from both, it is fatal to neither.

REFERENCES

Adby, P.R., and Dempster, M.A.H. *Introduction to Optimization Methods.* London: Chapman and Hall, 1974.

Caddy, G.R., and Lovibond, S.H. Self-regulation and discriminated aversive conditioning in the modification of alcoholics' drinking behavior. *Behavior Therapy* 7:223-230, 1976.

Campbell, D.T., and Stanley, J.C. *Experimental and Quasi-experimental Designs for Research.* Chicago: Rand McNally and Company, 1963.

Chaney, E.F.; O'Leary, M.R.; and Marlatt, G.A. Skill training with alcoholics. *J Consult Clin Psychol* 46:1092-1104, 1978.

Cox, W.M., and Klinger, E. A motivational model of alcohol use. *J Abnorm Psychol* 97:168-180, 1988.

Fisher, R.A. *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd, 1925.

Fuller, R.K.; Branchey, L.; Brightwell, D.R.; Derman, R.M.; Emrick, C.D.; Iber, F.L.; James, K.E.; Lacoursiere, R.B.; Lee, K.K.; Lowenstam, I.; Maany, I.; Neiderhiser, D.; Nocks, J.J.; and Shaw, S. Disulfiram treatment of alcoholism: A Veterans Administration cooperative study. *JAMA* 256:1449-1455, 1986.

Gilbert, F.S., and Maxwell, P.J. Predicting attendance at follow-up evaluations in alcohol treatment outcome research. *J Stud Alcohol* 48:569-573, 1987.

Howard, K.I.; Krause, M.S.; and Orlinsky, D.E. The attrition dilemma: Towards a new strategy for psychotherapy research. *J Consult Clin Psychol* 54:106-110, 1986.

Hull, C.L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943.

Kalton, G. *Compensation for Missing Survey Data.* Ann Arbor: Survey Research Center, 1983.

Kaplan, A. *The Conduct of Inquiry.* New York: Intext, 1964.

Lovibond, S.H. Use of behavior modification in the reduction of alcohol-related road accidents. In: Thompson, T., and Dockens, W.S., eds. *Applications of Behavior Modification.* New York: Academic Press, 1975. pp. 399-406.

Lovibond, S.H., and Caddy, G.R. Discriminated aversive control in the moderation of alcoholics' drinking behavior. *Behavior Therapy* 1:437-444, 1970.

Madow, W.G., ed. *Incomplete Data in Sample Surveys, Vols. I-III.* New York: Academic Press, 1983.

Mackenzie, A.; Funderbunk, F.R.; Allen, R.P.; and Stefan, R.L. The characteristics of alcoholics frequently lost to follow-up. *J Stud Alcohol* 48:119-123, 1987.

McCrady, B.; Longabaugh, R.; Fink, E.; Stout, R.; Beattie, M.; and Ruggieri-Authelet. Cost effectiveness of alcoholism treatment in partial hospital versus inpatient setting after brief inpatient treatment: 12-month outcome. *J Consult Clin Psychol* 54:708-713, 1986.

Moos, R.H., and Finney, J.W. The expanding scope of alcoholism treatment evaluation. *Am Psychol* 38:1036-1044, 1983.

Riley, D.M.; Sobell, L.C.; Leo, G.I.; Sobell, M.B.; and Klajner, F. Behavioral treatment of alcohol problems: A review and a comparison of behavioral and nonbehavioral studies. In: Cox, W.M., ed. *Treatment and Prevention of Alcohol Problems: A Resource Manual.* Orlando, FL: Academic Press, 1987.

Rosenthal, R. *Experimenter Effects in Behavioral Research.* New York: Irvington Press, 1976.

Sanchez-Craig, M.; Wilkinson, D.A.; and Walker, K. Theory and methods for secondary prevention of alcohol problems: A cognitive approach. In: Cox, W.M., ed. *Treatment and Prevention of Alcohol Problems: A Resource Manual.* Orlando, FL: Academic Press, 1987. pp. 287-331.

Skinner, B.F. *The Behavior of Organisms.* New York: Appleton-Century-Crofts, 1938.
Smith, M.L.; Glass, G.V.; and Miller, T.I. *The Benefits of Psychotherapy.* Baltimore: The Johns Hopkins University Press, 1980.
Sobell, L.C. Critique of alcoholism treatment evaluation. In: Marlatt, G.A., and Nathan, P.E., eds. *Behavioral Approaches to Alcoholism.* New Brunswick, NJ: Rutgers Center of Alcohol Studies, 1978. pp. 166-182.
Stephan, F.J., and McCarthy, P.J. *Sampling Opinions.* New York: Wiley, 1958.
Underwood, B.J. *Experimental Psychology.* New York: Appleton-Century-Crofts, 1949.

ACKNOWLEDGMENTS

This chapter is based, in large part, on a 1986 paper by Howard, Krause, and Orlinsky. This work was partially supported by research grant R01 MH42901 from the National Institute of Mental Health.

AUTHORS

Kenneth I. Howard, Ph.D.
Professor of Psychology
Department of Psychology
Northwestern University
Evanston, IL 60208

W. Miles Cox, Ph.D.
North Chicago Veterans Administration Medical Center
University of Health Sciences/The Chicago Medical School
North Chicago, IL 60064

Stephen M. Saunders, Ph.D.
Department of Psychology
Northwestern University
Evanston, IL 60208

Conceptualizing and Selecting Measures of Treatment Outcome: Implications for Drug Abuse Outcome Studies

Michael J. Lambert

INTRODUCTION

Those investigating treatment effects have been largely occupied with the nature of the treatment process, while ignoring or focusing on assessment of change as only an afterthought. As a consequence, thinking and theorizing about the assessment of change lags behind developments in treatment strategies themselves. Unfortunately, the development of effective treatments for drug abuse or any disorder is largely dependent on the accurate assessment of client change.

In this chapter, a brief history of assessment is presented, followed by a general strategy for organizing assessment strategies. The purpose of the strategy presented is to facilitate the general goals of empirical research on psychosocial and medical interventions—to increase the certainty and power of the methods we use to help people overcome their problems.

HISTORICAL OVERVIEW OF OUTCOME MEASUREMENT

Although measurement and quantification are central properties of empirical science, the earliest attempts at quantifying treatment gains lacked scientific rigor. Table 1 suggests several dimensions upon which assessments have varied during the relatively short history of studying outcome. The field has gradually moved from complete reliance on therapist ratings of gross/general improvement to the use of outcome indices of specific symptoms that are quantified from a variety of viewpoints, including the patient, outside observers, relatives, physiological indices, and environmental data such as employment records. The data generated from these

TABLE 1. *Developmental history of outcome assessment*

Change Rated by Therapist	——>	Multiple Sources
Ratings of Gross Change	——>	Specific Change/Multiple Technology
Theory Bound	——>	Practically Important
Change Is Unidirectional.	——>	For Better or Worse
Change Is Unidimensional.	——>	Change Is Multidimensional.
Changes Are Stable.	——>	Changes Are Unstable.

viewpoints are always subject to the limitations inherent in the methodology relied upon; none are "objective" or most authoritative, but they represent an improvement from previous measurement methods, which were difficult to replicate.

Attempts at evaluating psychotherapy have frequently reflected current in-vogue theoretical positions. Thus, early studies of psychotherapy applied devices that developed from Freudian dynamic psychology. Not at all uncommon was the use of projective methodologies, including the Rorschach Inkblot Test, the Thematic Apperception Test (TAT), Draw-a-Person, and Sentence-Completion methods. Problems with the psychometric qualities of these tests, their reliance on inference, and their derivation from a theoretical position based on the unconscious all resulted in their waning use as indices of outcome. Rarely today does one hear the virtues of such tests for outcome measurement. Changes in the quality of fantasy material as produced on the TAT and other projective tests simply do not convince most researchers and observers that significant improvement has occurred in the actual lives of patients.

Projective methodology gave way to assessment devices derived from other theories. Client-centered psychology, for example, concentrated on measures of perceived-self, ideal-self discrepancies based on the Q-sort technique. These and related measures of self-concept proved only slightly better than projective techniques. The disappearance of such devices as the sole index of improvement in outcome research and acceptance of the idea that the effects of psychotherapy should extend into the daily functioning of patients must be viewed as signs of progress.

These theoretically derived devices have been replaced by theoretical measures of factors such as adequate role performance, symptomatology, and the direct observation of target behaviors.

Clearly, a review of past and current practices for assessing psychotherapy outcome indicates that the field, while not without its problems, is maturing. Assessment procedures are becoming more complex and are also relying more heavily on standardized instruments that deal with specific kinds of change. Although there are many problems with current measurement methods, they hold considerable excitement, if not promise, for the future. But despite the current interest in outcome assessment, there is a long way to go before consensus will be reached on the type of yardstick to apply to the results of psychotherapy and behavior-change techniques with even homogeneous populations such as drug abusers. One has the impression that despite improvements in assessment practices, greater effort needs to be directed toward comprehensive assessments of change. Researchers are not prone to clarify the limits of their assessment practices. They do not discuss the general philosophy underlying their choices of instruments or the implications of such choices.

DIVERSITY IN OUTCOME MEASUREMENT

Although outcome research has divorced itself from the sole use of theoretically based, single measures of change, the result has been great divergence in the criteria used.

Froyd and Lambert (1989), after reviewing assessment practices in outcome studies published in 20 major journals between the years of 1983 and 1988, found no fewer than 1,430 distinct measures applied in 348 studies. The type, number, and quality of measures varied greatly across journals, disorders, and treatment methods. Some data from this literature review are presented in table 2.

TABLE 2. *Frequency and percentage of measures by content*

Content	Frequency	Percentage
Intrapersonal	1,058	74
Interpersonal	240	17
Social Role Performance	132	9
Total	1,430	

NOTE: Data base: 348 psychotherapy outcome studies published in 20 journals between 1983 and 1988.

Even with homogeneous samples and treatments, there is great diversity. For example, Ogles and Lambert (1989) studied the assessment practices used in controlled outcome studies of agoraphobia published between 1966 and 1988. The results of the review are shown in table 3.

Inspection of these data shows that over 135 separate measures were employed in the 170 studies analyzed. This is amazing diversity when one considers that the focus of treatment (agoraphobia), goals of treatment, and the intervention employed (mostly cognitive behavioral therapy) were very

TABLE 3. *Number of agoraphobia instruments used in 170 outcome studies by category of measure*

Category	Number
Fear and Anxiety Measures	27
Behavioral Measures	10
Depression	9
Mental Evaluation	8
Unstandardized Rating Scales	39
Physiological	6
Personality and General Symptoms	14
Others	22
Total	135

limited in scope. Although outcome research has divorced itself from the sole use of theoretically based, single measures of change, the result has been great divergence in the criteria used. In agoraphobia research, even in the category where one would expect consensus (measures of anxiety and panic), there is still a long way to go before researchers agree on measures (see table 4).

Studies reporting on the treatment outcome of drug-abusing clients are no exception to this rule. Wells and colleagues (1988), for example, reviewed the ways in which researchers had studied "drug use" in outcome studies. This review ignored the many other outcomes that could be and were assessed, such as employment or arrest records, and concentrated on measures that focused on **drug usage** only. Even here, there were five categories of use: measures of consumption, categorical classification, weighted indices of seriousness, composite indices of problem severity, and patterns of use. These five categories together had more than 25 distinct procedures and measures to assess drug use!

Needless to say, the diversity in measurement practices has not made it easy to sum across studies and draw sound conclusions about the effectiveness of specific treatments, interactions between treatment and client variables, and similar crucial questions that can never be answered by a single outcome study. The simplicity and lack of precision in early studies has given way

to such great diversity, and even chaos, that there is an obvious need for integration and organization.

TABLE 4. *Fear/agoraphobia/panic/anxiety questionnaires*

Instruments	Frequency*
Fear Questionnaire (1979 Version)	41
Fear Survey Schedule (1964 Version)	34
Fear Questionnaire (1970 Version)	7
Fear Survey Schedule	3
Agoraphobia Questionnaire	1
Fear of Negative Evaluation	2
State-Trait Anxiety	5
Phobic Anxiety and Avoidance Scales	88
Zung Anxiety Scale	12
Taylor Manifest Anxiety Scale	12
Hamilton Anxiety Rating Scale	14
Social Avoidance and Distress	2
Agoraphobia Inventory	3
Acute Panic Inventory	3
Phobia Rating Scale	1
Gurney Phobic Scale	3
Fear of Autonomic Sensations	1
Social Phobia/Agoraphobia Inventory	1
Burns Agoraphobia Questionnaire	1
IPAT Anxiety Scale	3
Phobic Survey Schedule	1
Rubin Fear Survey Schedule	1
Hillside Scale of Functional Capacity	1
Hillside Acute Panic Inventory	1

*Number of studies that used the scale

A PROPOSED CONCEPTUAL SCHEME

Because divergent processes are occurring in therapeutic change and people themselves embody divergent dimensions or phenomena, divergent methods of criterion measurement must be used to capture accurately the complexity of human functioning. If change is multidimensional and several instruments are needed to reflect this change, then what guiding principles might most profitably direct researchers in their choice of measures? To a great degree, researchers are bound by practical constraints. These constraints are likely to include time, money, and the needs and comfort of clients. A theoretically sound and comprehensive list of instruments must usually give

way to these practical considerations. Despite this, an ideal scheme may be presented for the purpose of giving direction to, and illuminating the limitation of, the final assessment package. Figure 1 presents such a scheme. It has evolved from earlier, simpler schemes that have been most typically based on dichotomies, such as the idea that outcome should represent measures of "dynamic" vs. "symptomatic" improvement; "internal" vs. "behavioral" changes; "source" vs. "surface" traits; and the like (Lambert et al. 1986).

Content	Technology	Source	Temporality
Intrapersonal	Evaluation	Self-Report	Single Measure
1	1	1	
Affect	2	2	
1	•	•	
•	Description	Trained Observers	Repeated Measure
•	1	1	
Behavior	2	2	
1	•	•	
2	•	•	
•	Observation	Relevant Other	Pattern Measure
•	1	1	
Cognition	2	2	
1	•	•	
2	•	•	
•	Status	Therapist Rating	
•	1	1	
Interpersonal	2	2	
1	•	•	
2	•	•	
•		Institutional	
•		1	
Social Role		2	
Performance		•	
1		•	
2			
•			

FIGURE 1. *Scheme for organizing and selecting outcome measures*

Most categorizations of outcome measures have been based on practical grounds, emphasizing the source providing the data, or have been derived directly from existing psychotherapy interventions. Researchers have had difficulty in finding a unified frame of reference from which to view outcome measures—one that gives due consideration to both theoretical and methodological concerns. Shall we consider outcome from the point of

view of source of data, type of rating scale, aim of assessment, type of patient problem, process producing the data, or some other point of view?

While researchers' choice of outcome measures may be determined by their theory of behavior change and mutual agreement with clients over desired outcomes, it is recommended that attention also be given to a pragmatic view of comprehensive outcome assessment. Such a conception places due emphasis on the fact that different **sources** of ratings correlate at modest levels but seem to provide separate information that is important to consider. It also recognizes that the data can be collected by various methods or **technologies**. These can take the form of behavioral observations, psychophysiological monitoring, judgments, descriptions, and the like. Each of these methods of data collection presents a different view of the change process, and perhaps each should be represented in outcome assessment. For the sake of convergent validity, at least two different technologies ought to be employed.

Another important aspect of assessment is the degree to which it reflects a single assessment of a person or multiple measures of behavior over time. Both Orlinsky (1988) and Turner (1988) have recommended measures that broaden the focus to patterns of behavior reflected from multiple time periods rather than the typical measure of time-limited functioning used in outcome measurement today. It is important to classify measures with regard to the degree to which they may represent a life pattern.

In addition to temporality, source, and technology, one must consider range of content. The focus of evaluation could be on varied content areas such as mood, symptoms (psychopathology), or self-concept; role performance, self-regulation, or self-control; and physical performance. It is recommended that content be divided into three categories: intrapersonal, interpersonal, and social-role performance.

The categories reflect the need to assess outcomes that range from those affecting mainly the individual and the symptoms that are an indication of disturbance to those that reflect the individual's intimate relationship with significant others and those that emphasize the individual's relationship and contribution to society. We have further subdivided intrapersonal content into affect, cognition, and behavior (including physiological responding). One defense for such a division is the tendency of different therapies to emphasize one or another of these aspects of people to the exclusion of others. A single outcome study, using a wide variety of assessment sources and multiple "methods" of assessment, can hardly hope to assess adequately improvement in all relevant content areas. The result of this dilemma is that progress will be slow, but without systematic efforts at understanding the interrelationship of sources, methods, content areas, and stability, progress will be even slower.

It is recommended that researchers attempt to assess change with full attention to content, technology, source, and temporality. When their efforts fall short of comprehensive assessment, the inherent limitations of this methodology should be noted.

There is some evidence that the dimensions mentioned can and do influence outcome assessment. A limited sample of this evidence will give the reader a sense of its nature. In the study of agoraphobia by Ogles and Lambert (1989), effect sizes showed remarkable differences as a possible function of a mixture of the above dimensions. These data are presented in table 5.

TABLE 5. *Overall effect size by frequently used scale*

Scale	Number of Treatments (n)	Mes	S Des
Phobic Anxiety/Avoidance	80	2.44	1.75
Global Assessment Scale	31	2.30	1.14
Self-Rating Severity	53	2.10	1.5
Fear Questionnaire	56	1.92	1.3
Anxiety During BAT	48	1.36	.84
BAT	67	1.18	.99
Depression Measures	74	.99	.70
Fear Survey Schedule	39	.96	.63
Heart Rate	21	.44	.56

KEY: Mes=mean effect size across n treatment groups.
S Des=standard deviation of effect sizes across n treatment groups.
BAT=Behavioral Avoidance Test

As can be seen, the size of treatment effects between pre- and post-time periods varies from a low of .44 for heart-rate change to a high of 2.44 for self-reported changes in phobic anxiety and avoidance.

Another interesting finding was extracted from data reported by Shapiro and Shapiro (1982). (See table 6.)

These researchers showed that the technology or methods of data collection that they found across studies showed significant correlations with treatment effects. Of special interest to our discussion is the finding that more reactive measures, such as those used in the Behavioral Avoidance Test, and

TABLE 6. *The relationship of outcome measure characteristics to outcome*[a]

Characteristics	Correlation of Effect Size by Characteristic	Outcome
Domain (Tractability)	.09**	More Tractable Larger Effect
Reactivity	.11**	More Reactive Larger Effect
Specificity	.07*	More Specific Larger Effect
Technology	-.11**	Softer Measures Larger Effect

*p=.05

**p=.01

[a]Based on the meta-analysis of Shapiro and Shapiro (1982)

KEY: n=1,828 effect sizes

softer measures, such as our evaluation category, are significantly associated with larger effect sizes. Researchers can anticipate that technology impacts outcome, but in just what ways is an empirical question that has rarely been answered.

Finally, data extracted from another meta-analytic review published by Miller and Berman (1983) suggest that treatment effects vary as a function of source and that the conclusions drawn about treatment effects may depend on whether the analysis is based on one source rather than another (see table 7).

While much additional evidence could be provided, the preceding is sufficient to suggest that the conclusions we draw from psychotherapy outcome studies depend to a large degree on various aspects of the measures used. It is a matter of great importance that we increase our understanding of the ways outcome assessment impacts the search for powerful, efficient treatments. Government agencies such as the National Institute on Drug Abuse must continue their leadership in the scientific enterprise that may ultimately allow us to reduce much human suffering. An important priority in this leadership is the assessment of change. A conceptual and practical guide for assessment of research efforts should

assist in the evaluation of effective treatments and the integration of the research being reported across this country and the world.

TABLE 7. *Cognitive behavioral therapies (CBT) using different sources of outcome*

Source of Measure	Number of Studies	Mean Effect Size	SD
	Compared to No Treatment Control		
Self Report	38	.86**	.54
Observer	24	.70	.41
Physiological	8	.03	.87
	Compared With Other Treatments		
Self Report	35	.21*	.37
Observer	23	.06	.43
Physiological	8	.03	.87

*p=.01

**p=.001, sig. from 0.

NOTE: When compared to no treatment, CBT appeared more effective on the basis of self-report and observer ratings but not physiological measures. When compared with other treatments, CBT appeared more effective only on the basis of self-report measures.

SOURCE: Abstracted from Miller and Berman 1983.

REFERENCES

Froyd, J., and Lambert, M.J. A Survey of Outcome Research Measures in Psychotherapy Research. Unpublished manuscript, Department of Psychology, Brigham Young University, Provo, UT, May 1989.

Lambert, M.J.; Shapiro, D.A.; and Bergin, A.E. The effectiveness of psychotherapy. In: Garfield, S.L., and Bergin, A.E., eds. *Handbook of Psychotherapy and Behavior Change* (3rd ed.). New York: John Wiley and Sons, 1986. pp. 157-211.

Miller, R.C., and Berman, J.S. The efficacy of cognitive behavior therapies: A quantitative review of the research evidence. *Psychol Bull* 94:39-53, 1983.

Ogles, B., and Lambert, M.J. A survey of outcome assessment in agoraphobia outcome research. Unpublished manuscript, Department of Psychology, Brigham Young University, Provo, UT, May 1989.

Orlinsky, D.E. Conceptualizing psychotherapy outcome. A panel presented at the 19th annual meeting of the Society for Psychotherapy Research, Santa Fe, NM, June 1988.

Shapiro, D.A., and Shapiro, D. Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychol Bull* 92:581-604, 1982.

Turner, R. Conceptualizing psychotherapy outcome. Presented at the 19th annual meeting of the Society for Psychotherapy Research, Santa Fe, NM, June 1988.

Wells, E.A.; Hawkins, J.D.; and Catalano, R.F. Choosing drug use measures for treatment outcome studies. I. The influence of measurement approach on treatment results. *Int J Addict* 23:851-873, 1988.

AUTHOR

Michael J. Lambert, Ph.D.
Department of Psychology
Brigham Young University
272 TLRB
Provo, UT 84602

Can A Technology Model of Psychotherapy Research Be Applied to Cocaine Abuse Treatment?

Kathleen Carroll and Bruce Rounsaville

INTRODUCTION

The technology model is one of the most powerful designs that can be used to evaluate psychotherapeutic treatments in randomized clinical trials. This model attempts to specify the treatment variable—psychotherapy—in a manner analogous to specification of a drug's formulation in pharmacological trials, that is, definition of treatment techniques in manuals as well as precise specification of treatment dose, delivery of treatments, nature of the subject sample, and therapists' experience and training. Through this specification, a technology model for psychotherapy research seeks to control extraneous variability in clinical trials in order to address differentiated questions concerning the conditions under which specified types of subjects will respond to particular types of treatments (Williams and Spitzer 1984).

The complex challenges in conceptualization and application of this model revolve around the central problem of internal vs. external validity. Internal validity, which requires precision and control in study design and execution of treatments so that alternative explanation of results may be ruled out, must be balanced against the representativeness of treatments evaluated and the generalizability of results. In a technology model of psychotherapy research, particular difficulty lies in the polarity between variables that must be controlled to reduce variation in psychotherapy in order to provide a valid technological evaluation of psychotherapeutic treatments and those variables that cannot or should not be controlled because of the inherent variability of psychotherapy (Docherty 1984; Waskow 1984).

The technological model of psychotherapy research reached one of its highest levels of development to date in the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Project (TDCRP) (Elkin et al. 1985; Elkin et al. 1988a; Elkin et al. 1988b). Here, extensive consideration was given to several issues: (1) specification, standardization, and dose of treatments; (2) nature of the control and reference conditions; (3) characterization of the subject population, which included use of standardized diagnostic instruments; (4) therapist characteristics and training; (5) monitoring of administration of the therapies through the course of the study; and (6) multidimensional assessment of outcome, which included measures of treatment specificity as well as use of independent evaluators who were blind to treatment received. Much of the technology pioneered in the development of the NIMH TDCRP has become standard for psychotherapy efficacy research in other psychiatric populations.

Can a technology model be applied to psychotherapy efficacy research in substance abuse? In this chapter, the authors highlight selected issues that pose methodological difficulties in the application of a technological model of psychotherapy research in clinical trials with ambulatory cocaine abusers. Those issues that are frequently encountered as potential threats to internal validity in this context are emphasized. These include problems related to the subject population, such as heterogeneity, selection, and attrition, as well as issues related to the delivery of treatments and the role of psychotherapy in treating cocaine abusers.

ISSUES RELATED TO THE SUBJECT POPULATION

Heterogeneity

The technology model requires specification and recruitment of a homogeneous subject sample to assure that study treatments are administered only to individuals who may expect to benefit from them and to minimize variability in outcome due to variability in subjects. Marked heterogeneity in a subject sample also introduces a number of nuisance variables, which, together with the small sample sizes characteristic of psychotherapy outcome research, may lead to nonequivalence between treated groups (Hsu 1989).

Although homogeneity in a study sample is clearly desirable, descriptions of treated cocaine abusers consistently emphasize their heterogeneity on a number of relevant dimensions, such as intensity and chronicity of cocaine use (Gawin and Kleber 1986; Schuster and Fischman 1985), concurrent use/dependence on other substances (Chitwood 1985; Siegel 1985), and presence of concurrent psychopathology (Gawin and Kleber 1986; Rounsaville et al., in press; Weiss et al. 1986; Weiss et al. 1988). Recent introduction of alternate forms of cocaine that are sold in smaller, less expensive quantities has resulted in wider availability of cocaine to individuals previously "protected" by cocaine's high cost. This may lead

to greater variability in level of psychosocial resources among treated cocaine abusers.

One strategy for controlling heterogeneity among cocaine abusers in outcome research is to exclude some subgroups of cocaine abusers. Here, one would attempt to balance representativeness and homogeneity in the subject sample, with the selection of subgroups to be excluded predicated on the nature of the treatments studied. For example, if one were studying a treatment developed specifically for cocaine abuse, it might be preferable to evaluate that treatment with cocaine abusers who do not use other psychoactive substances, as negative treatment outcome in a sample including polysubstance users might be difficult to interpret (as polysubstance users might represent a distinct subgroup with poor prognosis). On the other hand, exclusion of polysubstance abusers would lead to a more homogeneous but small and highly unrepresentative sample, as the majority of those whose principal drug of abuse is cocaine also manifest substantial concurrent use of other psychoactive substances (Rounsaville et al., in press). As some forms of substance abuse may represent a consequence of chronic cocaine use, e.g., the development of secondary alcohol dependence through the use of large amounts of alcohol to attenuate negative aspects of cocaine toxicity, exclusion of cocaine abusers with secondary substance abuse disorders would sharply reduce generalizability.

Similarly, subgroups of cocaine abusers with concurrent psychiatric disorders might be excluded from clinical trials because (1) a treatment for a primary substance-use disorder may be inappropriate in those cases where the psychiatric disorder is a major current problem or (2) considerable alteration in technique of a psychotherapy might be required to treat cocaine abusers with psychiatric disorders, and this could adversely affect the integrity of that therapy as a study treatment. Again, given the high rates of current psychiatric disorders in treatment-seeking substance abusers, exclusion of all cocaine abusers with coexistent psychiatric disorders would lead to an unrepresentative sample. To balance representativeness and homogeneity, investigators could exclude those cocaine abusers with psychiatric disorders for whom study treatments would be clearly contraindicated, such as schizophrenia, but include cocaine abusers with other types of psychiatric disorder and explore in the data analysis the effects of psychiatric morbidity on outcome, e.g., variation in outcome associated with different diagnostic categories or the significance of the principal/secondary distinction for the cocaine abuse and psychiatric disorder.

A different strategy for addressing heterogeneity is stratification, or assigning subjects to treatment conditions on the basis of patient characteristics thought to be related to outcome. Given the early stage of treatment research on cocaine abuse, however, few subject variables have as yet been related to outcome, and stratification may be premature. Nevertheless, stratification may be indicated when a particular treatment is

hypothesized to be powerfully related to a particular subject variable. For example, a study evaluating a form of psychotherapy whose rationale is amelioration of a principal psychiatric disorder that may be underlying the substance-use disorder might require stratification to assure adequate and balanced distribution of that disorder across treatment groups. Use of stratification on one variable does not guarantee balance between groups on other potential nuisance variables, and exploration of these in the data analysis may nonetheless be required.

A final strategy would be to use simple randomization to assign patients to study treatments and explore the relationship between subject variables where heterogeneity is marked (e.g., intensity and chronicity of use, route of administration, presence of concurrent psychopathology) and outcome in the data analysis, with post hoc analysis of subject variables strongly related to outcome as covariants. This approach is typical of early stages of outcome research in a new area and allows for exploration of the effectiveness of study treatments on the broad population with a particular disorder. Once effective treatments have been developed and evaluated through management trials, which may allow preliminary identification of subject variables related to outcome, "explanatory" trials, which require much more narrowly defined subject samples, may be undertaken (Sackett and Gent 1979).

Selection

Selection refers to those factors that operate to prevent some individuals from being included in an investigation. If these factors are substantial or systematic, an unrepresentative and highly skewed sample may result. Selection factors include (1) decisions made by subjects on whether or not they will participate in a trial (based on perceived risks/benefits of study treatments or the attractiveness of treatments offered); (2) decisions made by research staff on whether a potential subject is appropriate for the study (whether or not a potential subject meets inclusion/exclusion criteria); and (3) nature of the setting in which the research is conducted (particular types of patients may be drawn to different treatment settings, such as inpatient or outpatient clinics, or influenced by the reputation of a setting for providing certain types of treatments). In clinical trials, these factors may operate simultaneously, each affecting the composition of the sample that is ultimately included and studied.

Since selection cannot be avoided (the universe of treatment-seeking cocaine abusers cannot be included in any one study), the usual approach is to describe the effects of selection working in a particular investigation. For the first two types of selection (decisions made by either the subject or investigators regarding participation), this may be done by (a) documenting both the number and reasons for the exclusion of potential subjects, for example, the number and reason for those who do not fit inclusion/ exclusion criteria, and (b) describing the number of individuals who "reject"

the study and their putative reasons for not participating. In this way, important differences between subjects included in a study sample and those typically treated in the same site can be described.

In contrast, it is much more difficult to describe the influence of a particular research setting on individuals excluded from a trial because individuals who never approach the setting cannot be evaluated nor offered the opportunity of participating in the study. For this reason, it is advisable to describe, as extensively as possible, the characteristics of the sample finally included in a study and the setting in which they were treated. Generalization of findings to groups seen in comparable sites, for example, inner-city outpatient cocaine clinics, is likely to be more appropriate than generalization to dissimilar sites (a State-funded outpatient clinic vs. a private inpatient unit).

Finally, the type of selection that most influences determination of sample composition is one investigators can do very little about: Of the universe of cocaine abusers who may need treatment, only a very small fraction will actually seek treatment (Nash et al. 1965). Treatment seekers are therefore a very highly selected group of cocaine abusers, and the selection factors described above operate only to divide further this very highly selected subset. Results from any investigation conducted with treatment-seeking cocaine abusers may only reasonably be generalized to other treatment seekers.

Attrition

Postinclusion loss of subjects through attrition can threaten the validity of findings from treatment-outcome research in a number of ways: (1) Attrition operates to reduce sample size and therefore power; (2) Differential attrition between treatment cells may undermine randomization or stratification and produce nonequivalent groups; and (3) Differential attrition related to poor outcome may produce bias against treatments with lower attrition. Treatment-outcome research in substance abuse is fraught with high rates of attrition (Baekeland and Lundwall 1975; Craig 1985), and preliminary reports indicate similarly high rates of attrition among treated cocaine abusers (Anker and Crowley 1982; Gawin et al. 1989; O'Brien et al. 1988; Rawson et al. 1986).

The preferred strategy in approaching the problem of attrition is, of course, to prevent or reduce it. This may be attempted through a number of techniques. First, investigators should provide adequate preparation of subjects for study treatments through pretreatment information interviews. Preinclusion information or "socialization interviews" (Nash et al. 1965) provide detailed information on the nature of study treatments subjects might receive and explain the requirements of study participation, such as frequent completion of assessment instruments or random urinalysis. Such

interviews may reduce attrition by preventing unrealistic expectations of treatment or study procedures. Informational interviews may also clarify the conditions under which subjects may be withdrawn from the study, for example, failure to comply with treatments or marked clinical deterioration, so that subjects may be less likely to drop out after circumscribed episodes of cocaine use, as such subjects may assume they will be withdrawn from the study or admonished by study staff and not return to treatment. Second, rapid assignment to study treatments after application for treatment may reduce attrition among cocaine abusers, who often present for treatment in crisis and may be unable to persevere through a protracted pretreatment assessment process that delays actual treatment delivery. Third, a variety of supportive systems and procedures, such as flexibility in scheduling appointments and contacting subjects immediately after missed appointments, may also be helpful in reducing attrition. Availability of full-time research staff to handle subjects' questions and problems between regular appointments may be particularly important in those cases where study therapists are not part of the regular clinic staff and may not be easily reached by subjects outside regularly scheduled appointments.

Attrition can rarely be eliminated, but investigators may make use of a variety of reparative strategies in data analyses. These include techniques such as endpoint analyses, upper-bound estimation, or life-table analyses (Dodge 1985; Lasky 1962). Because no single strategy for treating attrition in data analyses is ideal (Howard et al. 1986), convergence of findings using multiple statistical methods may facilitate interpretation of results in trials where attrition is substantial. For example, if Treatment A is found to be superior to Treatment B on the basis of analyses of all subjects who completed treatment, endpoint analyses for all subjects who had at least moderate exposure to the treatments, and survival curves, one might be reasonably confident in assertions that Treatment A is more effective than Treatment B, as each method would utilize a slightly different sample and make different assumptions concerning the status of dropouts.

Finally, because rates of attrition in clinical trials of cocaine abuse treatments are so high, treatments that improve retention have marked advantage over those with high attrition. It is therefore critical that attrition be considered as an outcome variable and not merely a factor influencing missing data (Kazdin and Wilson 1978).

DELIVERY OF TREATMENTS

One of the more challenging aspects of the technology model for psychotherapy research is specification of the treatment variable. In this model, psychotherapies are defined and standardized in treatment manuals, and the conditions under which psychotherapies are administered, such as their duration, length, and frequency, are standardized as well. Furthermore, delivery of treatments in clinical trials traditionally reflects common clinical

practice and prevailing models of treatment. For example, during the period the NIMH TDCRP was developed, the prevailing models of treatment for depression were psychotherapy, pharmacotherapy, or their combination. This was reflected in the design of the NIMH TDCRP, which evaluated two forms of psychotherapy with a pharmacotherapy reference condition. In the treatment of opiate addiction, the prevailing model of treatment remains methadone maintenance; therefore, the major studies of psychotherapy for opiate addicts have evaluated psychotherapy as an adjunct to methadone maintenance (Rounsaville et al. 1983; Woody et al. 1983).

No prevailing model for the treatment of cocaine abuse has been established, and investigators are as yet relatively free to evaluate the delivery of psychotherapy in a variety of frameworks, for example, with or without pharmacotherapy or other supportive treatments. The adequacy of purely psychotherapeutic treatments for cocaine abuse, however, has not yet been demonstrated, and purely psychotherapeutic treatments may be insufficient for many cocaine abusers. Thus, investigators may choose to evaluate psychotherapies in the context of ongoing pharmacotherapy or other supports or to institute randomization to purely psychotherapeutic treatments only after stabilization, for example, via brief hospitalization. The impact of psychotherapy, however, may be diminished or difficult to demonstrate in the context of other treatments or after hospitalization, which may substantially affect outcome on their own.

While the lack of a prevailing treatment model does confer certain latitude in study design, several factors related to the delivery of treatments may influence internal validity, such as balancing treatments, subject expectancies, and control of adjunctive treatments.

Balancing Treatments

In order to maximize a study's feasibility and prevent bias, treatments to be compared must be comparable in terms of important parameters so that differences in outcome cannot be attributed to differences in the treatments' attractiveness, cost, or credibility. This may limit the kind of comparisons that may be offered in treatment-outcome research with cocaine abusers, as comparisons of widely different forms of treatment, such as inpatient vs. outpatient treatment, lengthy vs. brief psychotherapy, and family vs. individual treatment, may not be feasible. Furthermore, randomization to widely different (such as inpatient vs. outpatient) or unattractive treatments may not be accepted by substance abusers (Hall 1984). The issue of balance is particularly salient in selecting control conditions in psychotherapy research, as it is difficult to define psychotherapy control conditions that parallel "active psychotherapy" in intensity or credibility without becoming active psychotherapies themselves.

One strategy for offering balanced treatments in psychotherapy research is use of comparative designs, in which two active psychotherapies are compared. Comparative designs respond to several methodological and ethical questions associated with placebo or minimal treatment-control conditions, including differences in demand characteristics and credibility of the treatments as well as lack of control for nonspecific elements of the therapies (Basham 1986; Kazdin 1986; O'Leary and Borkovec 1978; Parloff 1986; Wilkins 1984; Wilkins 1986). If differential treatment effects are not found in a comparative design, however, it may be difficult to determine whether study treatments have been more effective than no treatment. Other strategies for offering balanced treatments are discussed elsewhere in this volume and include dismantling, component control, and parametric designs, in which a single intervention is varied in frequency between treatment conditions.

Expectancies

Subjects may have clear preferences or expectations for treatment, and this may influence their acceptance of study treatments and/or their response to treatments. Among cocaine abusers, who may leave treatment rather than persist with treatments they think are not credible, desirable, or effective, it is critical to attempt to prevent subjects' expectations from adversely affecting treatment retention or outcome. For example, if a subject is randomized to a treatment for which he/she has low expectations for success, the effectiveness of that treatment may be diminished because the patient may not comply with treatment or may be less willing to persist with treatment if a positive response is not rapid. While cocaine abusers may or may not be able to articulate a clear preference for one form of psychotherapy over another, many have clear preferences for the modality in which treatment is delivered, such as inpatient vs. outpatient treatment or individual vs. group therapy, or the form of treatment they receive, such as medication or psychotherapy.

In designs where the impact of differential treatment expectations on outcome is likely to be relatively minor, assessment of subjects' expectations or treatment preferences may be an acceptable strategy. For example, if treatments being compared are structurally similar (as in a comparative study of two forms of psychotherapy or in component control designs), subjects' differential expectations for one form of treatment vs. another may be relatively small. In those cases, it may be adequate merely to measure subjects' expectations and then to explore the effect of the congruence of treatment expectations or preferences and actual treatments received and treatment outcome.

In contrast, if widely different treatments are compared, subjects' expectations for treatment type may be substantially different. Here, in addition to assessing expectations, it may be useful to offer combination

treatments so that the violation of subjects' strong preferences for one form of treatment or another would be less likely to lead to attrition or dissatisfaction with treatment. An example would be a 2x2 design in which all patients receive some form of psychotherapy (either an active or control condition) and some form of pharmacotherapy (either active medication or placebo). In this way, subjects expressing a strong preference for medication would probably accept the combination of another treatment (psychotherapy) with the desired treatment (medication).

Adjunctive Treatments

In psychotherapy efficacy research, adjunctive treatments, such as family therapy, vocational training, or self-help groups, are a threat to internal validity because (1) adjunctive treatments may reduce a subject's level of involvement or effort in study treatments and therefore adversely affect treatment integrity; (2) adjunctive treatments may be differentially sought by subjects in control conditions; and (3) adjunctive treatments may powerfully affect outcome independently or in their combination with study treatments. Consequently, the usual practice in psychotherapy research is to restrict adjunctive treatments during the course of the study or to exclude those individuals who require adjunctive treatments.

Prohibition of adjunctive treatments during clinical trials may be difficult in practice; hence, allowing participation in some types of adjunctive treatment may be an acceptable strategy, providing one can document the "dose" of such treatments in order to treat participation in adjunctive treatments as covariants or process variables in the data analysis. For example, because many investigators perceive self-help groups as providing crucial supports or as essential steps to successful outcome, they may be hesitant to restrict cocaine abusers from attending self-help meetings during the study protocol. In such cases, investigators might allow or encourage involvement with self-help groups, carefully documenting the number of such meetings attended. A danger is that study treatments may be overpowered by participation in self-help groups because self-help groups are likely to take place much more frequently than study treatments and may exert substantial influence on subjects. Consequently, some restriction may be warranted, such as asking subjects to refrain from participating in self-help groups during the first few weeks of a study, until a working alliance is established, or asking subjects to limit the number of meetings attended. Similarly, investigators who consider some contact with family members necessary might allow one or two family meetings, restricting the content of these meetings to the family's questions about the patient's treatment and explication of the rationale of study treatments, with prohibition of any interventions that could be considered family or marital therapy.

While limitation of adjunctive treatments during the course of a trial is relatively straightforward, prohibition of treatment during followup is more

problematic. First, cocaine abuse, like other substance-use disorders, tends to be a chronic disorder characterized by frequent relapse, and few subjects can be expected to be completely "cured" during the course of a time-limited research protocol. Consequently, prohibition of treatment during followup may forbid critically needed treatments for some subjects. Second, given high rates of attrition in substance-abuse treatment, a measure of a short-term study treatment's success might be the extent to which it facilitates induction into longer term treatments; therefore, asking subjects to leave treatment at the end of a research protocol would be counter-therapeutic. Moreover, restriction of treatment-seeking during followup is impossible to enforce.

Rather than forbidding further treatment during followup, investigators might allow subjects completing a clinical trial to continue in treatment, or get another type of treatment if they wish to do so, with any treatments received during followup monitored as closely as possible. This more naturalistic strategy may result in a number of potential problems in interpreting results from followup in that subjects may receive treatments that are extensive or widely different from those received as their study treatments. For example, subjects randomized to a discussion control for 12 weeks may receive several months of intensive psychotherapy during the followup period, or subjects randomized to a psychotherapy-only condition might receive active medication during followup. One method of handling such occurrences would be to consider followup data for groups of subjects who did and did not receive any treatment during followup in separate statistical analyses; another method would be to analyze followup data with respect to treatments actually received during followup, regardless of study treatments received. Such strategies, however, may reduce power considerably.

CONCLUSIONS

We have described selected issues to be considered in application of a technological model of psychotherapy-efficacy research to cocaine abusers. These methodological considerations relate to potential threats to internal validity posed by either (1) the nature of the subject population or (2) the delivery of treatments within the constraints of prevailing models of treatment. Each of these issues illustrates, to some degree, the dichotomy between internal and external validity. In providing strategies for addressing these problems, the authors have emphasized those that seek to preserve internal validity in the context of the practical problems raised in treating cocaine abusers in clinical trials. While these strategies may be imperfect, this model remains the most promising for evaluating the role of psycho-therapy in the treatment of cocaine abuse and for identifying the most promising psychotherapeutic strategies for the treatment of cocaine abuse.

Why not use this model? Major arguments against a technology model in randomized clinical trials for substance abusers relate to problems of external validity, such as (1) highly selected and unrepresentative samples and (2) the danger that the degree of control and specification of treatment required in this model so alters the nature of psychotherapeutic treatments that their evaluation is no longer meaningful. With regard to the first objection, selection cannot be avoided, whether or not a technology model is used, because the universe of treatment-seeking cocaine abusers cannot be included in any one study, clinic, or interval of time. With regard to the second objection, William Blake noted that "Art and science cannot exist but in minutely organized particulars"—meaningful evaluation of treatments is impossible without their specification. Specification of psychotherapy technique permits evaluation, modification, and improvement of specific interventions and the illumination of the processes of behavior change. It is therefore more likely to further the development of effective treatments than wholesale application of approaches where little attempt is made to define or control technique and interventions and all therapeutic activity is allowed to vary.

The strength of a technology model, which is its emphasis on internal validity, usually exacts some sacrifice in external validity. The alternatives, which are designs that emphasize generalizability, for example, naturalistic studies or quasi-experimental designs, do not approach the power of randomized clinical trials in ruling out alternative explanations of results, nor do they provide sensitive tests of a treatment's effectiveness.

Does the sacrifice in external validity invalidate the process? The studies that have provided the most convincing evidence to date of the efficacy of psychotherapy in the treatment of substance abuse have been those using a technology model in clinical trials, particularly those done by Woody and colleagues (1983) and Rounsaville and colleagues (1983). These investigations included features such as manual-guided therapies, carefully described subject populations, use of experienced therapists, multidimensional assessment of outcome, and independent raters who were blind to treatment condition. The advantages of the technology model—greater control, exportability, and replicability—were clearly demonstrated in these studies, which did much to confirm the effectiveness of psychotherapy for opiate addicts and to identify subpopulations most likely to obtain benefits.

Psychotherapy efficacy research in cocaine-abuse treatment is a nascent field, which at this time is more appropriately concerned with the empirical identification of effective psychotherapies than with their generalizability. Optimum psychotherapeutic strategies for treating cocaine abuse will be identified only through their definition, careful specification, and evaluation within this method. Once effective psychotherapeutic treatments are identified through well-controlled trials, research efforts can shift to evaluation of their generality by systematic replication with other samples,

BEST COPY AVAILABLE

in other settings, and with variation in therapy technique (Kazdin 1980). Use of the technology model to evaluate treatments of cocaine abuse remains more on the order of researchers with a problem well served by a model rather than "people with a method who are looking for a problem to use it on" (Stevens 1957).

REFERENCES

Anker, A.L., and Crowley, T.J. Use of contingency contracts in specialty clinics for cocaine abuse. In: Harris, L.S., ed. *Problems of Drug Dependence, 1981.* National Institute on Drug Abuse Research Monograph 41. Rockville, MD: the Institute, 1982. pp. 452-459.

Baekeland, F., and Lundwall, L. Dropping out of treatment: A critical review. *Psychol Bull* 82:738-775, 1975.

Basham, R.B. Scientific and practical advantages of comparative design in psychotherapy outcome research. *J Consult Clin Psychol* 54:88-94, 1986.

Chitwood, D.D. Patterns and consequences of cocaine use. In: Kozel, N.J., and Adams, E.H., eds. *Cocaine Use in America: Epidemiologic and Clinical Perspectives.* National Institute on Drug Abuse Research Monograph 61. Rockville, MD: the Institute, 1985. pp. 111-129.

Craig, R.J. Reducing the treatment drop out rate in drug abuse programs. *J Subst Abuse Treat* 2:209-219, 1985.

Docherty, J.P. Implications of the technological model of psychotherapy. In: Williams, J.B.W., and Spitzer, R.L., eds. *Psychotherapy Research: Where Are We and Where Should We Go?* New York: Guilford, 1984.

Dodge, Y. *Analysis of Experiments With Missing Data.* New York: Wiley, 1985.

Elkin, I.; Parloff, M.B.; Hadley, S.W.; and Autry, J.H. NIMH treatment of depression collaborative research program: Background and research plan. *Arch Gen Psychiatry* 42:305-316, 1985.

Elkin, I.; Pilkonis, P.A.; Docherty, J.P.; and Sotsky, S.M. Conceptual and methodological issues in comparative studies of psychotherapy and pharmacotherapy, I: Active ingredients and mechanisms of change. *Am J Psychiatry* 145:909-917, 1988a.

Elkin, I.; Pilkonis, P.A.; Docherty, J.P.; and Sotsky, S.M. Conceptual and methodological issues in comparative studies of psychotherapy and pharmacotherapy, II: Nature and timing of treatment effects. *Am J Psychiatry* 145:1070-1076, 1988b.

Gawin, F.H., and Kleber, H.D. Abstinence symptomatology and psychiatric diagnosis in cocaine abusers. *Arch Gen Psychiatry* 43:107-113, 1986.

Gawin, F.H.; Kleber, H.D.; Byck, R.; Rounsaville, B.J.; Kosten, T.R.; Jatlow, P.I.; and Morgan, C.B. Desipramine facilitation of initial cocaine abstinence. *Arch Gen Psychiatry* 46:117-121, 1989.

Hall, S.M. Clinical trials in drug treatment: Methodology. In: Tims, F.M., and Ludford, J.P., eds. *Drug Abuse Treatment Evaluation: Strategies, Progress, and Prospects.* National Institute on Drug Abuse Research Monograph 51. Rockville, MD: the Institute, 1984. pp. 88-105.

Howard, K.I.; Krause, M.S.; and Orlinsky, D.E. The attrition dilemma: Toward a new strategy for psychotherapy research. *J Consult Clin Psychol* 54:106-110, 1986.

Hsu, L.M. Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *J Consult Clin Psychol* 57:131-137, 1989.

Kazdin, A.E. *Research Design in Clinical Psychology.* New York: Harper & Row, 1980.

Kazdin, A.E. Comparative outcome studies of psychotherapy: Methodological issues and strategies. *J Consult Clin Psychol* 54:95-105, 1986.

Kazdin, A.E., and Wilson, G.T. Criteria for evaluating psychotherapy. *Arch Gen Psychiatry* 35:407-416, 1978.

Lasky, J.J. The problem of sample attrition in controlled treatment trials. *J Nerv Ment Dis* 135:332-337, 1962.

Nash, E.; Howehn-Saric, R.; Battle, C.; Stone, A.; Imber, S.; and Frank, J. Systematic preparation of patients for short-term psychotherapy. *J Nerv Ment Dis* 140:374-383, 1965.

O'Brien, C.P.; Childress, A.R.; Arndt, I.O.; McLellan, A.T.; Woody, G.E.; and Maany, I. Pharmacological and behavioral treatments of cocaine dependence: Controlled studies. *J Clin Psychiatry* 49:17-22, 1988.

O'Leary, K.D., and Borkovec, T.D. Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *Am Psychol* 33:821-830, 1978.

Parloff, M.B. Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems. *J Consult Clin Psychol* 54:79-87, 1986.

Rawson, R.A.; Obert, J.L.; McCann, M.J.; and Mann, A.J. Cocaine treatment outcome: Cocaine use following inpatient, outpatient, and no treatment. In: Harris, L.S., ed. *Problems of Drug Dependence, 1985.* National Institute on Drug Abuse Research Monograph 67. Rockville, MD: the Institute, 1986.

Rounsaville, B.J.; Foley, S.; Carroll, K.M.; Budde, D.; Prusoff, B.A.; and Gawin, F.H. Psychiatric diagnosis of treatment seeking cocaine abusers. *Arch Gen Psychiatry*, in press.

Rounsaville, B.J.; Glazer, W.; Wilber, C.H.; Weissman, M.M.; and Kleber, H.D. Short-term interpersonal psychotherapy in methadone maintained opiate addicts. *Arch Gen Psychiatry* 40:629-636, 1983.

Sackett, D.L., and Gent, M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 301:1410-1412, 1979.

Schuster, C.R., and Fischman, M.W. Characteristics of humans volunteering for a cocaine research project. In: Kozel, N.J., and Adams, E.H., eds. *Cocaine Use in America: Epidemiologic and Clinical Perspectives.* National Institute on Drug Abuse Research Monograph 61. Rockville, MD: the Institute, 1985. pp. 158-170.

Siegel, R.K. New patterns of cocaine use: Changing doses and routes. In: Kozel, N.J., and Adams, E.H., eds. *Cocaine Use in America: Epidemiologic and Clinical Perspectives.* National Institute on Drug Abuse Research Monograph 61. Rockville, MD: the Institute, 1985. pp. 204-220.

Stevens, S.S. On the psychophysical law. *Psychol Rev* 64:153-181, 1957.

Waskow, I.E. Specification of the technique variable in the NIMH Treatment of Depression Collaborative Research Program. In: Williams, J.B.W., and Spitzer, R.L., eds. *Psychotherapy Research: Where Are We and Where Should We Go?* New York: Guilford, 1984.

Weiss, R.D.; Mirin, S.M.; Griffin, M.L.; and Michael, J.L. Psychopathology in cocaine abusers: Changing trends. *J Nerv Ment Dis* 176:719-725, 1984.

Weiss, R.D; Mirin, S.M.; Michael, J.L.; and Sollugub, A.C. Psychopathology in chronic cocaine abusers. *Am J Drug Alcohol Abuse* 12:17-29, 1986.

Wilkins, W. Psychotherapy: The powerful placebo. *J Consult Clin Psychol* 52:570-573, 1984.

Wilkins, W. Placebo problems in psychotherapy research: Social-psychological alternative to chemotherapy concepts. *Am Psychol* 241:551-556, 1986.

Williams, J.B.W., and Spitzer, R.L., eds. *Psychotherapy Research: Where Are We and Where Should We Go?* New York: Guilford, 1984.

Woody, G.E.; Luborsky, L.; McLellan, A.T.; O'Brien, C.P.; Beck, A.T.; Blaine, J.; Herman, I.; and Hole, A. Psychotherapy for opiate addicts: Does it help? *Arch Gen Psychiatry* 40:639-645, 1983.

ACKNOWLEDGMENT

Preparation of this manuscript was supported by National Institute on Drug Abuse grant DA 04299.

AUTHORS

Kathleen Carroll, Ph.D.
Bruce Rounsaville, M.D.

Substance Abuse Treatment Unit
Yale University School of Medicine
27 Sylvan Avenue
New Haven, CT 06519

Methodology: What Are the Design Issues Involved in the Defined Research Priorities?

Larry E. Beutler

INTRODUCTION

The breadth of research priorities defined here necessitates maximal design flexibility. Ideally, treatment studies of drug abuse would employ a large number of actual clinicians, treatments identical to those used in clinical practice, a representative sampling of treatment settings, and actual patients identical to that group to which hypotheses are to be generalized. The limited resources available to most investigations emphasize the desirability of continuity of funding and integration among various research programs.

If done within the context of a sound set of programmatic objectives, sequential studies, some of which lack a degree of methodological rigor, may be more important than a large number of better controlled studies that are not integrated with one another. Indeed, a program of research may systematically include studies that vary in experimental control and research design precisely so that overall programmatic objectives are accomplished.

Campbell and Stanley (1963) describe three types of research designs, varying in degree of methodological control but each with some specific advantages. For example, preexperimental research, based upon case studies and clinical observations, has little ability to confirm hypotheses and lacks the experimental controls that allow it to stand convincingly on its own merit. Preexperimental research is useful, however, both to develop hypotheses for more systematic study and to illustrate the clinical meaningfulness of findings based upon more rigorous designs. Similarly, quasi-experimental research designs, based upon naturally derived groups and nonsystematically controlled treatments, lack the basis for causal inference. Nonetheless, they have the advantage of efficiency in the accumulation of large samples and are appropriate for the study of psychotherapy as it

occurs naturally and as it is delivered in settings and situations that are not amenable to experimental control.

The various methodological issues that must be addressed in individual research investigations in order to enhance the likelihood of valid findings can be summarized under the general headings of: (1) subject selection, (2) treatment definition, (3) therapist selection, (4) measuring treatment processes, (5) measuring treatment outcomes, and (6) issues related to data analysis.

SUBJECT SELECTION

Three issues are deserving of specific concern in the selection of subjects for psychotherapy research in drug abuse. These include the use of analog designs, recruitment, and the issue of sample characteristics.

Analog Research

Analog research represents an effort to translate the principles of clinical practice into situations that allow maximal control and thereby circumvent some of the ethical and practical problems of working directly in clinical settings. There are two types of analogs that have been applied to research in psychotherapy, but only one of these is ordinarily appropriate to studies of drug abuse. Subject analogs are probably inappropriately applied to drug abuse problems because of qualitative differences that are likely to exist between the analog population and the drug population by virtue of the etiology and consequences of drug abuse.

On the other hand, method analogs attempt to extract the relevant principles of a given clinical intervention and to apply them in a laboratory setting under controlled conditions and free from aspects of the treatment and context that might obscure the effects of the specific intervention. Thus, the analogy is between the clinical treatment and the laboratory intervention. Such studies can be conducted on clinical samples and provide an efficient way of testing theoretical principles.

In one sense, all controlled research embodies a method analog. Once one begins to take systematic measurements before and/or after treatment, observe or control therapist behavior, and alter treatment conditions or contexts, there is some departure from the usual method of delivering clinical services. Recognizing this, Kazdin (1986) has emphasized that there are degrees of similarity between analog studies and their clinical referents: the closer the correspondence, the more representative the study and the more likely the findings are valid and replicable.

Analog research is most valuable when studies based upon dismantling and parametric strategies are conducted (Borkovec, this volume). It is at this

point that specific interventions are being dismantled and methods of altering the pattern and means of delivery are being sought. A limited but favorable argument can be launched on behalf of such investigations as applied to constructive (systematically adding and testing components) and process-oriented strategies.

Subject Recruitment

Howard (this volume) has observed that the reduction of potential subject pools, by self-selection, by the imposition of stringent entry criteria, and by postassignment attrition, guarantees that samples will lack representativeness. The more specific the questions being addressed, the more narrow the limits between subject inclusion and exclusion, and the larger the number of potential subjects required in order to derive a meaningful sample, the less representative is the sample of the population of drug abusers to which results are to be generalized.

It cannot be assumed that those who refuse assignment, those who are selected out, and those who drop out after treatment assignment represent the same population as those who remain in treatment. Hence, there is a need to seek predictors of attrition in historical and current events as well as among patient-intake characteristics and treatment-process variables. Researchers should be encouraged to monitor closely patients who are self-selected out of controlled research protocols, those who are screened out prior to assignment, and those who drop out after assignment. Post hoc and systematic comparisons of these groups may begin to unravel the complex limits of generalization introduced by the processes that threaten random assignment.

The sources from which patient samples are recruited further strain the assumptions of randomization. Direct requests for volunteers are likely to result in a sample composed largely of the curious and the financially unfortunate. Requests to other agencies and health professionals for referrals are likely to receive little response. Unless the investigator has access to and control of a large clinic, it is unlikely that patients in the number needed for group comparison studies can be obtained through usual referral sources. Under these circumstances, the investigator has two options available. He/she either relies on designs based upon individual cases or the inclusion of naturalistic procedures.

Single-case research designs are valuable, especially for the initial development of hypotheses and instruments and for the assessment of process-to-outcome linkages (Kazdin 1986). Even with these designs, however, representativeness must be assured through replication and systematic variation. Time-series analyses, multiple baselines, and the application of statistical procedures appropriate to these designs all assist in defining and generalizing the findings.

All investigations should give attention to power functions. As a rule of thumb, power should be sufficient to detect changes of approximately 10 percent (the average proportion of variance attributable to treatments [Shapiro and Shapiro 1982]) or symptom alleviation in approximately 10 percent of the sample at least 50 percent of the time.

Naturalistic studies ordinarily allow the use of larger samples than do controlled studies. In these naturalistic studies, the focus is not only on the nature of patients, but also on the nature and representativeness of the treatment settings and formats that are sought and applied outside the controlled research program. Multisetting research is imperative for obtaining a wide and representative variety of naturally occurring treatment programs.

Sample Characteristics

A third area of methodological concern in sample selection pertains to the desirability of describing sample characteristics as fully as possible. This concern includes the need to define the clinical status of the sample, a process that is assisted by the use of standardized diagnostic procedures. Similarly, demographic data and information about marital status, support systems, previous treatment, and living conditions can provide important information for defining the limits of treatment efficacy.

Effective research programs include both common or core data and data that are specific to each research project. To the degree possible, investigators should utilize instruments that have been used in other investigations with similar populations. Common data sources allow comparison of patient characteristics and responses across studies.

While samples should be homogeneous with respect to drugs of choice and diagnosis, the crosscutting nature of psychotherapy procedures also suggests the desirability of exploring the interaction of treatment characteristics and other patient dimensions. The selection of appropriate dimensions should take into account both clinical experience and theoretical formulations. Potential interactive variables include polydrug use, patient coping styles, severity of problems, subjective discomfort levels, support networks, cognitive functioning, resistance patterns, development levels, and comorbidity.

TREATMENT DEFINITION

Manualized therapies have furthered the aims of standardizing treatments. Most treatments for drug abuse, however, draw procedures from a wide range of psychosocial interventions and diverse theoretical models. Specifying the procedures making up treatment packages is necessary for the important task of assessing the influence of treatments. Beyond the importance of demonstrating treatment integrity and efficacy, application of manualized treatment packages has limited value. Each patient is unique,

and each clinician attempts to account for this uniqueness when developing treatment plans. To approximate the way that clinical practice actually works, investigators should remember that the ultimate translation of research findings to clinical practice will be made when each treatment selected is unique to the needs of each patient. To do this, research must ultimately move past the inspection of global treatment packages to consider the specific components comprising these packages.

Dismantling and constructive-research strategies will be required in the context of outcome and process-oriented research in order to determine the specific relationships among procedures, patients, therapists, and responses. Due attention should be given, within controlled research programs, to the roles of treatment duration, frequency, modality, method of patient preparation, level of intervention, mediating tasks and goals of treatment, breadth of treatment focus, directiveness, the use of both intratherapy and extratherapy session tasks, and nature of the treatment relationship.

For the field to advance, manuals are needed by which to study and direct the selection of psychosocial procedures extracted from several therapeutic theories. Such manuals would detail the processes of clinical decision-making and may begin to provide the bases for integration in the field.

THERAPIST SELECTION

An effective research program will ensure that adequate data are gathered on all therapists to subsequently assess the influence of their own demographic backgrounds, social support systems, beliefs, values, personality styles, theoretical philosophies, expectations, preferences, and goals.

Aside from this background and characterological data, therapist training is critical for most studies of psychotherapy efficacy. In controlled research designs, training therapists to criterion-defined levels of skill and competence is needed in order to assure consistent and powerful tests of research hypotheses. There remain important research questions as to how effectively therapists can be trained to perform procedures from several different, often conceptually inconsistent, treatment programs. Within-therapist comparisons as well as between-therapist comparisons are needed to determine the limits of therapists' abilities to use a diversity of procedures.

Research on the educational and learning process must develop hand in hand with studies of treatment efficacy in order to ensure that the principles discovered in outcome research are translatable to clinical practice. It does little good to demonstrate that treatments differ in effectiveness for different types of drug-abusing patients if clinicians are not also assured that they can learn to perform the procedures discriminatively and competently. Variations in treatment procedure should accommodate to the variations of

duration, formats, and modalities of treatment that characterize clinical practice.

MEASURING TREATMENT PROCESSES

Issues related to understanding and controlling psychotherapeutic processes are also relevant to the need to translate research knowledge to clinical practice. To be clinically meaningful, the problems selected for study must have clinical appeal, the treatments used must bear a close relationship to actual treatments, and clinical processes and outcomes must be seen as relevant to clinical practice. Instruments must be sensitive to clinically meaningful dimensions, supply information of value to the clinician, and not be so intrusive as to compromise the integrity of treatment.

When approaching the task of assessing treatment processes, several considerations must be raised. Some of these relate to the need for general measures that cut across treatments, some relate to the nature of patient/clinician relationships, and some relate to the need for treatment-specific measures. Treatments vary in nature and specificity. They are somatic; they are social; they are psychotherapeutic; and they are none of these things exclusively. There is some advantage to selecting variables and measures that yield sufficiently broad dimensions as to be relevant to all these domains of experience. Omnibus measures of treatment process allow easy comparison among treatment variations. This comparison is especially important in those instances when outcome data are expected to be interpretable or valid only much later in the research program and when one must rely on inferred relationships between treatment processes and outcomes in order to assess treatment effects within an acceptable timeframe.

Careful record keeping is the key to assessing treatment processes in broadly based and wide-ranging treatments. Meaningful process measures can be derived by recording the nature of treatment as completely as possible and doing so in quantifiable terms. The amount and frequency of contact with the treating clinician, the amount and type of medication, the nature of the treatment setting, the duration of treatment, the nature and frequency of collateral contacts, the characteristics of the clinician, and patient satisfaction with how closely treatment matches initial expectations are all variables that can be gathered relatively independently of any specific treatment program through careful records. While there are few specific tools for such wide-ranging assessment, it is of some value to develop program-specific data sets that at least roughly correspond across programs and studies. The process of developing archival data sets can be enhanced by taking advantage of computerized packages for systematically gathering patient data. Computer information and record-keeping systems have been developed for hospital-based treatment programs (Crawford

et al. 1974), community mental health systems (Hedlund et al. 1977; Sherman 1981), and multisite State-run programs (Laska 1981).

In addition to recording informational data about the nature of treatment assignment, there is also some value in assessing both specific and general aspects of the clinician/therapist relationship. A variety of instruments are available to assess both general, non-theory-bound dimensions and specific processes that are valued by therapies that vary in level of intervention and theory specificity. Among the general process measures, some rely on lexical aspects of speech, while others assess aspects of therapeutic flow, disruption, and topic initiation. Among the more theory-specific measures, those relating to patient/therapist relationship qualities are most prominent. These include various means of assessing the therapeutic alliance and patient emotional experiencing. A comprehensive review of these various instruments and their uses in assessing aspects of the treatment relationship has been compiled by Greenberg and Pinsof (1986).

In addition to the foregoing, there are a variety of aspects of the treatment process for which measures are yet to be developed. These areas of needed instrument development include the assessment of theory-specific mediating goals of treatment, the assessment of treatment-task resolution, the assessment of changes in task-relevant arousal and focus, and easily administered measures of states of resistance or reactance.

Measures of patient and therapist congruity are especially needed for effective work to proceed on the nature of the patient/therapist match. Effective and reliable means for assessing congruity of initial and change expectations and congruity of interpersonal needs, beliefs, and values are especially needed.

MEASURING TREATMENT OUTCOME

The assessment of treatment outcome is a particularly difficult and multi-faceted problem. The estimate of outcome achieved is a function of (1) the source of rating, (2) the nature of the instruments used, (3) the method of estimating benefits, and (4) the point in time at which measures are made.

Source of Rating

The meanings of outcome estimates in mental health research are closely related to the source of the ratings, i.e., the role of the person making the ratings. Different raters approach the task of assessing treatment outcome from very different perspectives and rely on different types of data. Treating clinicians may evaluate benefit from a theoretical perspective, while patients may view change in terms of symptom change and family members may evaluate improvement on the basis of interpersonal conflict or patient compliance. Patients, therapists, and external observers may be restricted in

the realm of behavior to which they have access and also have investments that distort or bias their ratings. Hence, while there is often correspondence among data sources, when differences occur, these differences are likely to reflect biasing or limiting factors, such as rater expectations and hopes, as well as the nature of the most easily obtained observations.

Because of differences among rating sources, it is advisable that both clinicians and researchers actively seek to base estimates of progress and benefit upon several sources rather than simply rely on those obtained through a single perspective. Only by combining and comparing divergent viewpoints is one likely to obtain a clear picture of the effects of treatment. Lambert and colleagues (1983) have provided a detailed review of the variety of instruments available from these several sources as well as recommendations for their use.

Another measurement problem arises with respect to the need for individually tailored change estimates. In spite of their overlap with global estimates of change, there is a need for exploration and development of individualized measures that capture both treatment-relevant and patient-specific patterns of change. There are two procedures for constructing such individualized measures. The first of these is to employ N=1 study designs. This procedure is valuable because it allows the inspection of relatively rapid and clinically meaningful changes in outcome as a function of the many individualized decisions made by a clinician. The procedure also overcomes some of the problems generated by the differences in instrument reactivity and lack of conceptual similarity when one attempts to apply individualized measures in group designs.

The second method for assessing individualized change is to combine some specifically developed or selected instruments with a relatively standard set of "core" instruments. The core instruments promise to make the findings comparable to those of other studies both within and across research programs, while the individualized change assessment promises to make the data meaningful on an individual level. When a core battery is developed, recommended measures should cover aspects of symptoms, patient satisfaction, and social functioning. More specific assessments of change may then extend these areas of assessment by differentiating between symptomatic and conflictual changes and adding individualized components tailored to the areas of functioning considered specific within a given treatment model.

Estimating Benefits

Outcome criteria have always represented a difficult problem for treatment outcome research. This difficulty is largely traceable to disagreements among investigators about what constitutes meaningful dimensions on which

change is likely to occur, the means for comparing outcomes across studies, and the clinical significance of the benefits observed.

The question of what dimensions of change are relevant is compounded by the joint observations that patients, therapists, and external observers frequently disagree about relevance, and even when defined, the criteria may change over time in treatment. The use of a core assessment battery, combined with some specifically selected measures based upon clinician preference and patient-defined targets for change is important to cover the range of outcomes expected.

Responding to the concern that the different targets and scaling devices used by different studies prevented direct comparability of results, Smith and Glass (1977) pioneered the use of meta-analysis in treatment outcome. Effect Size (E.S.) estimates were expressed in terms of standard deviations such that an E.S. of 1.0 indicated that, on a particular change measure, the average patient in the treated group improved more than did 68 percent of patients in the untreated condition.

While meta-analyses have been criticized on a number of grounds, their position in treatment-outcome research is now firmly established (Shapiro and Shapiro 1982). The use of standardized scores to express change in treatment groups relative to control conditions led to the development of methods for making research findings translatable into estimates of clinical as opposed to merely statistical significance, a longstanding bone of contention between researchers and clinicians. Increasing numbers of procedures are evolving for assessing the clinical meaningfulness of outcomes (Jacobson et al. 1986; Rosenthal and Rubin 1985).

Temporal Considerations

Treatments that characteristically produce few differences when measured at the end of a course of treatment often produce very different findings when patients are evaluated some months later. The speed of treatment as well as the longevity of treatment benefits should be considered in assessing outcomes. Similarly, there is an increasingly apparent need for long-term followup studies and relapse prevention programs. Investigators should carefully monitor characteristics of patients and environments that might be predictive of speed, strength, and duration of changes and that might differentiate treatments.

ANALYSIS

A variety of decisions related to the nature of the analysis procedure used in any research program have wide-ranging implications for how meaningful and reliable the findings will be. These issues cannot be divorced completely from those addressed in the foregoing pages. For example, the

sample characteristics, the selection of instruments, and the method used for deriving change estimates may set limits on the nature of the analysis procedures available. Beyond these concerns, however, when the investigator selects statistical methods, he/she must address questions of how to compensate for (1) missing data, (2) familywise error rates, and (3) the presence of multicolinearity among data sets. These questions are inextractable from decisions about the use of specific statistical procedures.

Missing Data

Every research program is plagued by the problem of missing data, incurred because of either subject dropout or noncompliance. To the points that have already been made about tracking dropouts and evaluating patient, contextual, and social contributors to such phenomena over time, the author can add only the emphasis upon a priori definitions about what constitutes an adequate "dose" of treatment on which to make reasonable comparisons. Largely, this decision is based upon the nature of the treatment and relies on the clinical wisdom of the investigators to define at what point a minimal amount of treatment has been delivered so that a treatment effect logically could be expected. While defining such a criterion of "treatment" does not eliminate the need to assess the contributors to dropout prior to this defined point, it does minimize the problem of defining what constitutes the "treated group."

Beyond the definition of treatment dose, a number of methods have been proposed for handling missing data and treatment dropouts that occur after the criterion amount of treatment. The critical concern in these proposals is to ensure that differential compliance and dropout rates do not bias the findings. Assuming that disgruntled patients are the most likely to be noncompliant and to leave treatment prematurely, it is logical to assume that those who remain in a treatment program are likely to have better response than those who drop out. If dropout and compliance rates differ among comparison groups, the conditions with the highest noncompliance and dropout rates are likely to show the most positive gains if only the end-of-treatment scores are compared.

A relatively conservative method for compensating for differential dropout rates when assessing treatment outcomes is to utilize variations of endpoint analysis. In one of these methods, each subject's last evaluated status is used at all subsequent assessment points as a replacement for missing values. Hence, endpoint analysis is always based upon the number of subjects that complete the minimal treatment dose on the assumption that if a patient improved beyond this point, it was probably not a direct result of treatment.

A more difficult problem arises when data points are missing because of noncompliance rather than premature termination. When patients/subjects

fail to provide complete information on a given questionnaire or fail to comply with one of several instruments or temporal evaluation points, the question of representativeness becomes critical once again. Basically, three avenues are open to the researcher in compensating for the potential biasing effects of this selective compliance. The first and least acceptable method is to drop the subject altogether from the analyses of those instruments and/or occasions in which noncompliance occurred. Unfortunately, this procedure can result in severely curtailed data sets and eliminates the use of many of the procedures required to combine and collapse instruments in order to reduce redundancy.

A second method for compensating for noncompliance on dependent measures is to replace missing data points with the average response of other patient/subjects assigned to the same comparison group. This procedure has the advantage of conservatively reclaiming data for uncooperative subjects and reducing the amount of attrition based upon noncompliance but has the disadvantage of assuming a degree of similarity among patients/subjects that frequently is unjustifiable.

The third and preferable method for reclaiming missing data because of noncompliance is to estimate the missing response from prior and following responses made by the same patient/subject. The easiest procedure is simply to substitute for the missing value a mean of the prior two or three values or a mean of the preceding and following values when such data are available. While easy to apply, this procedure does not work when a patient/subject refuses a specific test but takes others in the series or when there is a systematic but nonlinear response to the dependent variable across time.

A much more sophisticated procedure than simple averaging has been suggested by Welch and coworkers (1983). This procedure entails the use of regression analyses (probit and tobit analyses) to predict missing values based upon the patterns reflected in all prior and/or subsequent data of either a given subject or, when necessary, groups of subjects.

In this latter case, identifying the score to be inserted in place of missing values would involve first deriving a predictive algorithm based upon the relationship between responses to the targeted instrument and responses to instruments on which the subject was compliant, utilizing the entire cell sample. This would be followed by predicting each missing score derived from applying the algorithm derived from the entire sample to the specific scores of the noncompliant subject. The advantages of regression procedures such as this include both the ability to replace relatively large data sets that are missing and the ability to account for nonlinear patterns of response that may characterize different samples.

Compensating for Familywise Error Rates

When analysis is based on a finite sample and consists of several independent analyses, the probability of a Type I error is inflated multiplicatively by a factor consisting of the number of independent analyses undertaken.

The compensations for familywise error rates are typically to utilize a supraordinate analysis to precede more specific contrasts or to apply a correction for the number of contrasts undertaken. The first of these methods is often cumbersome, and the alternative is the application of a correction based upon the number of contrasts or analyses to be performed.

Research programs should be designed to economize the number of analyses used. Multiple regression analyses, multiple analyses of variance, and path or sequence analyses should be preferred over simpler univariate analyses whenever possible.

The regression/discontinuity design and analysis (Trochim 1984) is especially promising as a global procedure for analyzing data from naturalistic studies. This design compensates for problems arising when random assignment of treatment is not possible. The procedure is ordinarily applied as a pretest to posttest design wherein subjects are divided into groups on the basis of a pretest variable. Subsequently, posttest scores of these treatment groups are compared by both slope and regression analyses; that is, differences between treatments are revealed by both mean changes and changes in pretreatment to posttreatment slopes accrued between groups.

Additionally, analyses designed for testing specific a priori defined relationships and patterns, e.g., Lissrell and planned comparisons, will add power to the analysis without sacrificing expected error rates.

Compensating for Multicolinearity

A problem related to familywise error rates is that of data redundancy. When several instruments are utilized in a given study, there is a high probability that much of what they measure will overlap. This is especially true given the observation made earlier that most instruments assessing change measure a common quality of general or global improvement. The concern can be extended to the issue of independent and control measures, too. That is, if several measures are utilized for assessing an organismic quality of the patient to use as a predictor of change, even though the concepts may be distinct, it is likely that the measured variables will intercorrelate with one another. To the degree that scores from instruments within either dependent or independent variable sets are intercorrelated, the results may be inflated or deflated.

Multicolinearity can be addressed quite simply by combining redundant instruments or scores. The most efficient method for accomplishing this is to perform a principal components analysis on the set of dependent-variable-change scores (and a separate one on the control and independent measures if several are used). From this initial analysis, composite scores can be derived to reflect both common variances and the specific targets of change defined in the variable set. Weighted combinations based upon eigenvalues to reflect the factors extracted is the most sophisticated procedure, but one may, more simply, extract index scores from the various instruments based upon the size of factor loadings and the relative unique contributions made by these instruments to estimated change.

Of course, the use of factorial procedures in the manner proposed here encounters problems if based upon relatively small samples. Diaconis and Efron (1983) have described the use of high-intensive, computer-generated data sets to compensate for sample-size concerns. These procedures, e.g., bootstrapping, are especially designed to establish the reliability of small-sample findings, relying on only the assumption of sample representativeness.

CONCLUSIONS

In order to address the complexity of our clinical methods, programmatic research as well as individual investigations on drug abuse treatment are needed. Only in planned programmatic efforts will research be able to tease apart the various components and interactions among components that contribute to treatment efficacy. Moreover, only in programmatically derived sequences of studies can we hope to define the variables that are important to therapeutic efforts and achieve the large samples of patients that are needed statistically to assess the relevance of our postulations.

Research proceeds most efficiently by following theory. While the usual theories that guide clinical research are based upon the etiology of behavior, there is also a significant need for the development of models of treatment selection and decisionmaking that combine psychosocial interventions from different theoretical systems. Only by implementing a systematic program of research based upon a rational set of dimensions and principles of influence, however, can we hope to sort out the processes that make our treatments maximally successful.

REFERENCES

Campbell, D.T., and Stanley, J.C. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally & Co., 1963.

Crawford, J.L.; Morgan, D.W.; and Gianturco, D.T., eds. *Progress in Mental Health Information Systems: Computer Applications.* Cambridge, MA: Ballinger, 1974.

Diaconis, P., and Efron, B. Computer-intensive methods in statistics. *Sci Am* 248:116-130, 1983.

Greenberg, L.S., and Pinsof, W.M. *The Psychotherapeutic Process: A Research Handbook*. New York: Guilford, 1986.

Hedlund, J.L.; Sletten, I.W.; Evenson, R.C.; Altman, H.; and Cho, D.W. Automated psychiatric information systems: A critical review of Missouri's Standard System of Psychiatry (SSOP). *Journal of Operational Psychiatry* 8:5-26, 1977.

Jacobson, N.S.; Follette, W.C.; and Pagel, M. Predicting who will benefit from behavioral marital therapy. *J Consult Clin Psychol* 54:518-522, 1986.

Kazdin, A.E. The evaluation of psychotherapy: Research design and methodology. In: Garfield, S.L., and Bergin, A.E., eds. *Handbook of Psychotherapy and Behavior Change*, 3rd ed. New York: John Wiley, 1986. pp. 23-68.

Lambert, M.J.; Christensen, E.R.; and DeJulio, S.S., eds. *The Assessment of Psychotherapy Outcome*. New York: John Wiley, 1983.

Laska, E.M. Developments in computerization of the psychiatric record. In Siegal, C., and Krupnick-Fischer, S., eds. *Psychiatric Records in Mental Health Care*. New York: Brunner/Mazel, 1981. pp. 271-296.

Rosenthal, R., and Rubin, D.B. Statistical analysis: Summarizing evidence versus establishing facts. *Psychol Bull* 97:527-539, 1985.

Shapiro, D.A., and Shapiro, D. Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychol Bull* 92:581-604, 1982.

Sherman, P. A computerized CMHC clinical and management system: Saga of a "mini" success. *Behavioral Research Methods and Instrumentation* 13:445-453, 1981.

Smith, M.L., and Glass, C.V. Meta-analysis of psychotherapy outcome studies. *Am Psychol* 32:752-760, 1977.

Trochim, M.K. *Research design for program evaluation: The regression-discontinuity design*. Beverly Hills, CA: Sage Publications, 1984.

Welch, W.P.; Frank, R.G.; and Costello, A.J. Missing data in psychiatric research: A solution. *Psychol Bull* 94:177-180, 1983.

AUTHOR

Larry E. Beutler, Ph.D.
Professor, Psychiatry and Psychology
Chief of Psychology Programs
Director of Clinical Research
Director, Depression and Sleep Disorders Clinic
Department of Psychiatry
University of Arizona College of Medicine
Tucson, AZ 85724

Research Priorities and Future Directions

Research Priorities for Psychotherapy and Counseling in the Treatment of Drug Abuse: The Psychotherapy Research Perspective

Perry London

INTRODUCTION

Like Michael Lambert, I am less a therapy researcher than a consumer and wishful integrator of psychotherapy research (Lambert, this volume). My effort, therefore, will be to try to summarize the conference so far and set the framework for discussion by laying out what I believe the various speakers have taught us these past 2 days. Or, to borrow Tom McLellan's remark, I shall "torture the material" that has so far been presented to provide an invitational platform to your doing so in the discussion of research priorities (McLellan 1989).

This has been a remarkably stimulating and exciting conference. While it seems unlikely that Jack Blaine and Lisa Simon Onken actually rehearsed the papers with the people who presented them, they did a splendid job of orchestrating and prefacing them with remarks that are salient to all the papers subsequently presented (Blaine 1989; Onken 1989).

Their main thoughts emphasized the ubiquity of the use of counseling and psychotherapy in the treatment of drug abuse and the insufficiency of our understanding of whether and how these methods work. Most of the research on treatment of drug abuse has been drug treatment—yet, in at least 97 percent of the detoxification programs extant and 99 percent of the methadone treatment programs, most of what is actually done is some form of counseling or psychotherapy. It behooves us, therefore, to examine these much-used methods carefully to decide whether they are being done well, could be done better, or should be replaced by something more promising for the successful treatment of drug abuse.

THE GOALS OF DRUG ABUSE TREATMENT

In this connection, several speakers have made clear that the problem of treating drug abuse does not present the same Gordian knot that meta-analysis painfully cut through in evaluating most common forms of psychotherapy research. It is both easier and more difficult.

It is easier in the sense that drug abuse presents us with a clearer starting goal than do many psychotherapy problems—total abstinence. In another sense, however, this is a harder problem because the clients of drug abuse treatment programs have problems of motivation, distress, and objectives different from those we are accustomed to in most psychotherapy research.

For most kinds of psychotherapy, people seek treatment that corresponds directly to the kinds of pain they are suffering: If I am depressed, I want relief from depression; if I am anxious, I want freedom from anxiety. With drugs, that is very much not the case. People do not object to alcohol because they dislike being drunk. They do not abjure opiates or stimulants because they dislike being stoned. They object to them because of trouble in their lives resulting from the fact that they use those substances with pleasure and gratification. The problem of treating drug abuse conditions, accordingly, is more oblique with respect to the main symptoms and therefore makes the clients harder to treat.

Drug abuse is also a more complicated problem than those of "garden variety" psychotherapy, so to speak, because it is the "spillover effects" of the problem that chiefly bother society, even though substance use itself may be trouble enough. Those spillover effects are, on the face of it, matters that may be more directly addressed by counseling in the areas that Tom McLellan has mentioned, such as medical, legal, and family matters, than by what we commonly think of as either psychodynamic or behavioral therapeutic postures (McLellan 1989).

CONTENT STUDIES AND METHOD STUDIES

This conference has heard about some excellent empirical studies from scholars who are treating drug abuse with counseling and psychotherapy and about some excellent conceptual and methodological studies (and advice) from scholars who have been profoundly concerned with general psychotherapy research for many years. In that respect, it is noteworthy that a conference with this depth of inquiry and sophistication of method could not have taken place two decades ago, even on general psychotherapy, let alone on its specialized use in the treatment of drug abuse.

CONTENT STUDIES

Let us first look at the empirical studies of George Woody, Kathleen Carroll, and Paula Kleinman, this volume, which have been verified in their clinical observations and implications by Robert Millman's and Tom McLellan's presentations. What do they teach us? I think there are four general conclusions we can draw from them.

First, psychotherapy helps—perhaps most with patients who have other serious pathology in addition to drug problems. But it is not clear which of the different types of therapy tested is more effective: supportive-expressive therapy, interpersonal therapy, cognitive-behavior therapy, or relapse prevention therapy. Nor is it clear that there is a differential benefit derived from one or the other that makes it work better on problems of opiate or stimulant use.

Second, the variance among psychotherapists is large regardless of the kind of treatment they do—so large that psychotherapist differences may overwhelm and obscure technique differences.

Third, drug counselors are very important to the clinical treatment of drug abuse—perhaps especially for patients who do not have other severe psychopathology. I use the term "drug counselor" here, as opposed to "psychotherapist," to mean people who give advice and counsel on specific aspects of clients' lives. For drug abusers whose other psychological problems are not severe, drug counselors appear as effective as psychotherapists. But the exact and optimum roles of either counseling or psychotherapy functions are unclear, as are exact distinctions between them.

Fourth and finally, everyone who has presented at this conference has iterated, almost as a point of faith (since there are no data on the matter), that a 12-step program is effective for drug abusers (Woody, this volume; Carroll, this volume; Kleinman, this volume; Millman 1989; McLellan 1989).

METHOD STUDIES

The conceptual and methodological scholars also had four main points, each with an argument based in empirical study of a kind quite advanced over what we might have heard at conferences on psychotherapy research a dozen years ago.

First, Paul Crits-Christoph (this volume) told us that psychotherapist variance is even greater than George Woody (this volume) thinks. It is so important, he believes, that you should take several precautions in the design of process or outcome studies: (1) Use as many therapists as possible; (2) Always test for therapist effects; (3) If you choose to ignore

therapist effects, which you do at your peril, be very cautious in interpreting results; and (4) Try to avoid therapist effects by careful selection of therapists at the outset (Crits-Christoph, this volume).

Tom Borkovec (this volume) speaks not so much to **therapist** variance as to **therapy** variance. He says we should be doing in this area what has been done, successfully, in the analysis and evaluation of psychotherapy research hitherto: Looking at the **components** of the treatment packages we deliver and trying, in general, to set up parametric designs for research. Tom McLellan emphatically agrees (McLellan 1989). In addition, says Borkovec, we need to make a more pronounced effort to focus the direction of research by tying our empirical inquiries to theoretical concerns (Borkovec, this volume).

Kenneth Howard's (this volume) paper addresses neither therapist variance nor therapy variance, but **sample** variance. He argues that research should attend to sampling with close, even excruciating care, to **exclusion, inclusion,** and **attrition** criteria, and to what they mean to generalization. He also advocates the study of "naturalistic" (treatment-as-usual) treatment conditions.

He illustrates this message powerfully by describing a clinical study in which more than 6,000 people volunteered as subjects for alcoholism-treatment research and only 600 people were accepted. Some were rejected because they lived too far away; some because they had not had a drink recently enough; and some because they refused to accept the wording of the study's informed consent form. With such huge blanket exclusions, asks Howard, what exactly do you think we will learn about treatment that can be generalized from this sample to a meaningful population of alcohol abusers (Howard, this volume)?

Finally, Michael Lambert speaks not to therapist variance, therapy variance, or sample variance, but to the fact that if you wish to measure **outcome**, you must attend carefully to **outcome measures**. He astonished all of us by demonstrating the prolixity of extant measures and the huge variance across them in different studies. He makes a compelling argument for the careful organization and measurement of outcome and for how much the meaning of outcome results depends on the care with which the technology of measurement is managed in the first place when a study is designed (Lambert, this volume).

All four of these methodological considerations are very important to anyone who wants to work in this area.

NEEDED CONTENT RESEARCH: DRUG COUNSELORS AND 12-STEP PROGRAMS

In addition to the research design and method issues the speakers explicated, two content areas that need study were both implicit and explicit in much of the group's discussions and should probably be important aspects of research that NIDA should promote on the psychological treatment of drug abuse. The first is the issue of **drug counselors**, and the second is the issue of **12-step programs** and their equivalents.

Studying Drug Counselors

With respect to drug counselors, Robert Millman (1989) summed up a situation that George Woody had discussed previously and with which everyone clinically involved in this area seems to agree. Methadone drug counselors in particular, he stated, are in positions of low-skill expectations, low pay, and low prestige. At the same time, they are on the treatment firing line: A program's success or failure is attributed most directly to them. These people burn out in 6 to 12 months (Millman 1989). But even though drug counselors may be poorly prepared and poorly qualified for their work, George Woody has some evidence that their work is very important and sometimes effective. His research indicates that drug counseling has a function for some patients quite independent of that of psychotherapy (Woody, this volume), and the clinical notes that Tom McLellan has laid out agree thoroughly with it (McLellan 1989). If this is true, then both from a cost/benefit and a conceptual point of view, it may be terribly important for research to address specifically such questions as:

1. What role does drug counseling play in the treatment of drug abuse?

2. To what extent can it be distinguished and separated from psychotherapy? (Woody has discussed this to some degree already.)

3. Who should be doing which? Can the same people do both?

4. What are the mechanisms involved?

Studying 12-Step Programs

The second important content area needing study is the 12-step programs. Like everyone in the mental health business, I have been hearing for years about Alcoholics Anonymous (AA), about the 12 steps, and about the lore of treating substance abuse, which says: First, stop. Then get psychotherapy, and use AA as the best single device for the first part of that program and perhaps for the rest of it as well. Virtually everyone at this conference who is in the clinical side of drug treatment confirms that lore.

Robert Millman went so far as to say that he used to think it was invalid or unrealistic, but now he is "converted" (Millman 1989).

It may well be true that AA is the most effective treatment program extant for both alcohol and drug abuse. There is no doubt that it is the largest such program in the world. And now, having expanded its work to help the relatives and significant others of substance abusers, it is a still more important program than in the past.

But however convinced we are that AA is a wonderful program, we have no firm data on it. AA does not collect or publish any, and it does not support research on itself. The reasons for this may be perfectly good, but without hard data, the conviction among clinicians that AA is an ideal program is dubious because the drama sometimes involved in AA's effects may limit observation of their frequency and mislead us to overvalue them.

I have no doubt that AA cures some alcoholics and some drug abusers, but so does Billy Graham and so probably does the shrine at Lourdes. We must not dismiss or disparage the importance, the reality, or the stability of any of those cures. But we must, on the other hand, try to count them and account for them. The fact that some people are cured does not provide enough information by itself for us to develop public policy with respect to how those conditions and treatments could be best used in the hands of professionals.

COMPONENT PROGRAM RESEARCH

To develop a feasible policy, we need a secular research setting that will test methodologies like the 12-step programs in a context where data is formally collected and evaluated. I suspect that for programs like AA, there are at least three components that give them curative power. First, they involve a public commitment. We know from the social-influence literature, from Kurt Lewin's work through Jonathan Freedman's foot-in-the-door research, that public commitment has a relatively profound cementing effect on subsequent behavior. Second, the public commitment involved is made to relatively transcendent goals. There is reason to suspect, at least in the clinical lore, that this has a further reinforcing effect. Third, AA programs involve a social support network to reinforce behavior that helps people achieve declared goals and avoid relapse and to encourage recovery in the face of failure. There is already a sizable literature on stress and social support-network effects, and research in this area is expanding.

Thinking only in terms of those three variables—public commitment, transcendent goals, and social support networks—it should not be excessively difficult to set up what I would call a secular research program

in the treatment of drug abuse that could systematically evaluate the probable impact of clinical institutions such as AA.

REFERENCES

Blaine, J. Welcome. Presented at the National Institute on Drug Abuse Technical Review Meeting, Rockville, MD, May 18–19, 1989.

McClellan, A.T. Defining and measuring outcome in psychotherapy research studies. Presented at the National Institute on Drug Abuse Technical Review Meeting, Rockville, MD, May 18–19, 1989.

Millman, R. Psychotherapy with drug abusers: A clinical perspective. Presented at the National Institute on Drug Abuse Technical Review Meeting, Rockville, MD, May 18–19, 1989.

Onken, L.S. Opening remarks. Presented at the National Institute on Drug Abuse Technical Review Meeting, Rockville, MD, May 18–19, 1989.

AUTHOR

Perry London, Ph.D.
Professor and Dean
Graduate School of Applied and
Professional Psychology
Rutgers University
P.O. Box 819
Piscataway, NJ 08855

National Institute on Drug Abuse

Research

monograph series

While limited supplies last, single copies of the monographs may be obtained free of charge from the National Clearinghouse for Alcohol and Drug Information (NCADI). Please contact NCADI also for information about availability of coming issues and other publications of the National Institute on Drug Abuse relevant to drug abuse research.

Additional copies may be purchased from the U.S. Government Printing Office (GPO) and/or the National Technical Information Service (NTIS) as indicated. NTIS prices are for paper copy; add $3.00 handling charge for each order. Microfiche copies are also available from NTIS. Prices from either source are subject to change.

Addresses are:

NCADI
National Clearinghouse for Alcohol and Drug Information
P.O. Box 2345
Rockville, MD 20852

GPO
Superintendent of Documents
U.S. Government Printing Office
Washington, DC 20402

NTIS
National Technical Information Service
U.S. Department of Commerce
Springfield, VA 22161
(703) 487-4650

For information on availability of National Institute on Drug Abuse Research Monographs 1 through 24 (1975–1979) and others not listed, write to NIDA, Community and Professional Education Branch, Room 10A-54, 5600 Fishers Lane, Rockville, MD 20857.

25 BEHAVIORAL ANALYSIS AND TREATMENT OF SUBSTANCE ABUSE. Norman A. Krasnegor, Ph.D., ed.
GPO out of stock — NCADI out of stock
NTIS PB #80-112428/AS $31

26 THE BEHAVIORAL ASPECTS OF SMOKING. Norman A. Krasnegor, Ph.D., ed. (Reprint from 1979 Surgeon General's Report on Smoking and Health.)
GPO out of stock — NTIS PB #80-118755/AS $23

30 THEORIES ON DRUG ABUSE: SELECTED CONTEMPORARY PERSPECTIVES. Dan J. Lettieri, Ph.D.; Mollie Sayers; and Helen W. Pearson, eds.
GPO Stock #017-024-00997-1 — NCADI out of stock
Not available from NTIS

31 MARIJUANA RESEARCH FINDINGS: 1980. Robert C. Petersen, Ph.D., ed.
GPO out of stock — NTIS PB #80-215171/AS $31

32 GC/MS ASSAYS FOR ABUSED DRUGS IN BODY FLUIDS. Rodger L. Foltz, Ph.D.; Allison F. Fentiman, Jr., Ph.D.; and Ruth B. Foltz.
GPO out of stock — NCADI out of stock
NTIS PB #81-133746/AS $31

36 NEW APPROACHES TO TREATMENT OF CHRONIC PAIN: A REVIEW OF MULTIDISCIPLINARY PAIN CLINICS AND PAIN CENTERS. Lorenz K.Y. Ng, M.D., ed.
GPO out of stock — NCADI out of stock
NTIS PB #81-240913/AS $31

37 BEHAVIORAL PHARMACOLOGY OF HUMAN DRUG DEPENDENCE. Travis Thompson, Ph.D., and Chris E. Johanson, Ph.D., eds.
GPO out of stock — NCADI out of stock
NTIS PB #82-136961/AS $39

38 DRUG ABUSE AND THE AMERICAN ADOLESCENT. Dan J. Lettieri, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock — NCADI out of stock
NTIS PB #82-148198/AS $23

40 ADOLESCENT MARIJUANA ABUSERS AND THEIR FAMILIES. Herbert Hendin, M.D.; Ann Pollinger, Ph.D.; Richard Ulman, Ph.D.; and Arthur Carr, Ph.D., eds.
GPO out of stock — NCADI out of stock
NTIS PB #82-133117/AS $23

42 THE ANALYSIS OF CANNABINOIDS IN BIOLOGICAL FLUIDS. Richard L. Hawks, Ph.D., ed.
GPO out of stock — NTIS PB #83-136044/AS $23

44 MARIJUANA EFFECTS ON THE ENDOCRINE AND REPRODUCTIVE SYSTEMS. Monique C. Braude, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #85-150563/AS $23

45 CONTEMPORARY RESEARCH IN PAIN AND ANALGESIA, 1983. Roger M. Brown, Ph.D.; Theodore M. Pinkert, M.D., J.D.; and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #84-184670/AS $17

46 BEHAVIORAL INTERVENTION TECHNIQUES IN DRUG ABUSE TREATMENT. John Grabowski, Ph.D.; Maxine L. Stitzer, Ph.D.; and Jack E. Henningfield, Ph.D., eds.
GPO out of stock
NCADI out of stock
NTIS PB #84-184688/AS $23

47 PREVENTING ADOLESCENT DRUG ABUSE: INTERVENTION STRATEGIES. Thomas J. Glynn, Ph.D.; Carl G. Leukefeld, D.S.W.; and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO Stock #017-024-01180-1 $5.50
NCADI out of stock
NTIS PB #85-159663/AS $31

48 MEASUREMENT IN THE ANALYSIS AND TREATMENT OF SMOKING BEHAVIOR. John Grabowski, Ph.D., and Catherine Bell, M.S., eds.
GPO Stock #017-024-01181-9 $4.50
NCADI out of stock
NTIS PB #84-145184/AS $23

50 COCAINE: PHARMACOLOGY, EFFECTS, AND TREATMENT OF ABUSE. John Grabowski, Ph.D., ed.
GPO Stock #017-024-01214-9 $4
NTIS PB #85-150381/AS $23

51 DRUG ABUSE TREATMENT EVALUATION: STRATEGIES, PROGRESS, AND PROSPECTS. Frank M. Tims, Ph.D., ed.
GPO Stock #017-024-01218-1 $4.50
NTIS PB #85-150365/AS $23

52 TESTING DRUGS FOR PHYSICAL DEPENDENCE POTENTIAL AND ABUSE LIABILITY. Joseph V. Brady, Ph.D., and Scott E. Lukas, Ph.D., eds.
GPO Stock #017-024-01204-1 $4.25
NTIS PB #85-150373/AS $23

53 PHARMACOLOGICAL ADJUNCTS IN SMOKING CESSATION. John Grabowski, Ph.D., and Sharon M. Hall, Ph.D., eds.
GPO Stock #017-024-01266-1 $3.50
NCADI out of stock
NTIS PB #89-123186/AS $23

54 MECHANISMS OF TOLERANCE AND DEPENDENCE. Charles Wm. Sharp, Ph.D., ed.
GPO out of stock
NCADI out of stock
NTIS PB #89-103279/AS $39

55 PROBLEMS OF DRUG DEPENDENCE, 1984: PROCEEDINGS OF THE 46TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis. S. Harris, Ph. D., ed.
GPO out of stock
NCADI out of stock
NTIS PB #89-123194/AS $45

56 ETIOLOGY OF DRUG ABUSE: IMPLICATIONS FOR PREVENTION. Coryl LaRue Jones, Ph.D., and Robert J. Battjes, D.S.W., eds.
GPO Stock #017-024-01250-5 $6.50
NTIS PB #89-123160/AS $31

57 SELF-REPORT METHODS OF ESTIMATING DRUG USE: MEETING CURRENT CHALLENGES TO VALIDITY. Beatrice A. Rouse, Ph.D.; Nicholas J. Kozel, M.S.; and Louise G. Richards, Ph.D., eds.
GPO Stock #017-024-01246-7 $4.25
NTIS PB #88-248083/AS $23

58 PROGRESS IN THE DEVELOPMENT OF COST-EFFECTIVE TREATMENT FOR DRUG ABUSERS. Rebecca S. Ashery, D.S.W., ed.
GPO Stock #017-024-01247-5 $4.25
NTIS PB #89-125017/AS $23

59 CURRENT RESEARCH ON THE CONSEQUENCES OF MATERNAL DRUG ABUSE. Theodore M. Pinkert, M.D., J.D., ed.
GPO Stock #017-024-01249-1 $2.50
NTIS PB #89-125025/AS $23

60 PRENATAL DRUG EXPOSURE: KINETICS AND DYNAMICS. C. Nora Chiang, Ph.D., and Charles C. Lee, Ph.D., eds.
GPO Stock #017-024-01257-2 $3.50
NTIS PB #89-124564/AS $23

61 COCAINE USE IN AMERICA: EPIDEMIOLOGIC AND CLINICAL PERSPECTIVES. Nicholas J. Kozel, M.S., and Edgar H. Adams, M.S., eds.
GPO Stock #017-024-01258-1 $5
NTIS PB #89-131866/AS $31

62 NEUROSCIENCE METHODS IN DRUG ABUSE RESEARCH. Roger M. Brown, Ph.D., and David P. Friedman, Ph.D., eds.
GPO Stock #017-024-01260-2 $3.50
NCADI out of stock
NTIS PB #89-130660/AS $23

63 PREVENTION RESEARCH: DETERRING DRUG ABUSE AMONG CHILDREN AND ADOLESCENTS. Catherine S. Bell, M.S., and Robert J. Battjes, D.S.W., eds.
GPO Stock #017-024-01263-7 $5.50
NTIS PB #89-103287/AS $31

64 PHENCYCLIDINE: AN UPDATE. Doris H. Clouet, Ph.D., ed.
GPO Stock #017-024-01281-5 $6.50
NTIS PB #89-131858/AS $31

65 WOMEN AND DRUGS: A NEW ERA FOR RESEARCH. Barbara A. Ray, Ph.D., and Monique C. Braude, Ph.D., eds.
GPO Stock #017-024-01283-1 $3.50 NTIS PB #89-130637/AS $23

66 GENETIC AND BIOLOGICAL MARKERS IN DRUG ABUSE AND ALCOHOLISM. Monique C. Braude, Ph.D., and Helen M. Chao, Ph.D., eds.
GPO Stock #017-024-01291-2 $3.50 NCADI out of stock
NTIS PB #89-134423/AS $23

68 STRATEGIES FOR RESEARCH ON THE INTERACTIONS OF DRUGS OF ABUSE. Monique C. Braude, Ph.D., and Harold M. Ginzburg, M.D., J.D., eds.
GPO Stock #017-024-01296-3 $6.50 NCADI out of stock
NTIS PB #89-134936/AS $31

69 OPIOID PEPTIDES: MEDICINAL CHEMISTRY. Rao S. Rapaka, Ph.D.; Gene Barnett, Ph.D.; and Richard L. Hawks, Ph.D., eds.
GPO Stock #017-024-1297-1 $11 NTIS PB #89-158422/AS $39

70 OPIOID PEPTIDES: MOLECULAR PHARMACOLOGY, BIOSYNTHESIS, AND ANALYSIS. Rao S. Rapaka, Ph.D., and Richard L. Hawks, Ph.D., eds.
GPO Stock #017-024-1298-0 $12 NTIS PB #89-158430/AS $45

71 OPIATE RECEPTOR SUBTYPES AND BRAIN FUNCTION. Roger M. Brown, Ph.D.; Doris H. Clouet, Ph.D.; and David P. Friedman, Ph.D., eds.
GPO Stock #017-024-01303-0 $6 NTIS PB #89-151955/AS $31

72 RELAPSE AND RECOVERY IN DRUG ABUSE. Frank M. Tims, Ph.D., and Carl G. Leukefeld, D.S.W., eds.
GPO Stock #017-024-01302-1 $6 NTIS PB #89-151963/AS $31

73 URINE TESTING FOR DRUGS OF ABUSE. Richard L. Hawks, Ph.D., and C. Nora Chiang, Ph.D., eds.
GPO Stock #017-024-01313-7 $3.75 NTIS PB #89-151971/AS $23

74 NEUROBIOLOGY OF BEHAVIORAL CONTROL IN DRUG ABUSE. Stephen I. Szara, M.D., D.Sc., ed.
GPO Stock #017-024-1314-5 $3.75 NTIS PB #89-151989/AS $23

75 PROGRESS IN OPIOID RESEARCH. PROCEEDINGS OF THE 1986 INTERNATIONAL NARCOTICS RESEARCH CONFERENCE. John W. Holaday, Ph.D.; Ping-Yee Law, Ph.D.; and Albert Herz, M.D., eds.
GPO Stock #017-024-01315-3 $21 NCADI out of stock
Not available from NTIS

76 PROBLEMS OF DRUG DEPENDENCE, 1986. PROCEEDINGS OF THE 48TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01316-1 $16 NCADI out of stock
NTIS PB #88-208111/AS $53

77 ADOLESCENT DRUG ABUSE: ANALYSES OF TREATMENT RESEARCH. Elizabeth R. Rahdert, Ph.D., and John Grabowski, Ph.D., eds.
GPO Stock #017-024-01348-0 $4 NTIS PB #89-125488/AS $23

78 THE ROLE OF NEUROPLASTICITY IN THE RESPONSE TO DRUGS. David P. Friedman, Ph.D., and Doris H. Clouet, Ph.D., eds.
GPO Stock #017-024-01330-7 $6 NTIS PB #88-245683/AS $31

79 STRUCTURE-ACTIVITY RELATIONSHIPS OF THE CANNABINOIDS. Rao S. Rapaka, Ph.D., and Alexandros Makriyannis, Ph.D., eds.
GPO Stock #017-024-01331-5 $6 NTIS PB #89-109201/AS $31

80 NEEDLE SHARING AMONG INTRAVENOUS DRUG ABUSERS: NATIONAL AND INTERNATIONAL PERSPECTIVES. Robert J. Battjes, D.S.W., and Roy W. Pickens, Ph.D., eds.
GPO Stock #017-024-01345-5 $5.50 NTIS PB #88-236138/AS $31

81 PROBLEMS OF DRUG DEPENDENCE, 1987. PROCEEDINGS OF THE 49TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01354-4 $17 NTIS PB #89-109227/AS
Contact NTIS for price

82 OPIOIDS IN THE HIPPOCAMPUS. Jacqueline F. McGinty, Ph.D., and David P. Friedman, Ph.D., eds.
GPO Stock #017-024-01344-7 $4.25 NTIS PB #88-245691/AS $23

83 HEALTH HAZARDS OF NITRITE INHALANTS. Harry W. Haverkos, M.D., and John A. Dougherty, Ph.D., eds.
GPO Stock #017-024-01351-0 $3.25 NTIS PB #89-125496/AS $23

84 LEARNING FACTORS IN SUBSTANCE ABUSE. Barbara A. Ray, Ph.D., ed.
GPO Stock #017-024-01353-6 $6 NTIS PB #89-125504/AS $31

85 EPIDEMIOLOGY OF INHALANT ABUSE: AN UPDATE. Raquel A. Crider, Ph.D., and Beatrice A. Rouse, Ph.D., eds.
GPO Stock #017-024-01360-9 $5.50 NTIS PB #89-123178/AS $31

86 COMPULSORY TREATMENT OF DRUG ABUSE: RESEARCH AND CLINICAL PRACTICE. Carl G. Leukefeld, D.S.W., and Frank M. Tims, Ph.D., eds.
GPO Stock #017-024-01352-8 $7.50 NTIS PB #89-151997/AS $31

87 OPIOID PEPTIDES: AN UPDATE. Rao S. Rapaka, Ph.D., and Bhola N. Dhawan, M.D., eds.
GPO Stock #017-024-01366-8 $7 NTIS PB #89-158430/AS $45

88 MECHANISMS OF COCAINE ABUSE AND TOXICITY. Doris H. Clouet, Ph.D.; Khursheed Asghar, Ph.D.; and Roger M. Brown, Ph.D., eds.
GPO Stock #017-024-01359-5 $11 NTIS PB #89-125512/AS $39

89 BIOLOGICAL VULNERABILITY TO DRUG ABUSE. Roy W. Pickens, Ph.D., and Dace S. Svikis, B.A., eds.
GPO Stock #017-022-01054-2 $5 NTIS PB #89-125520/AS $23

90 PROBLEMS OF DRUG DEPENDENCE, 1988. PROCEEDINGS OF THE 50TH ANNUAL SCIENTIFIC MEETING. THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01362-5 $17

91 DRUGS IN THE WORKPLACE: RESEARCH AND EVALUATION DATA. Steven W. Gost, Ph.D., and J. Michael Walsh, Ph.D., eds.
GPO Stock #017-024-01384-6 $10 NTIS PB #90-147257/AS $39

92 TESTING FOR ABUSE LIABILITY OF DRUGS IN HUMANS. Marian W. Fischman, Ph.D., and Nancy K. Mello, Ph.D., eds.
GPO Stock #017-024-01379-0 $12 NTIS PB #90-148933/AS $45

93 AIDS AND INTRAVENOUS DRUG USE: FUTURE DIRECTIONS FOR COMMUNITY-BASED PREVENTION RESEARCH. C.G. Leukefeld, D.S.W.; R.J. Battjes, D.S.W.; and Z. Amsel, Ph.D., eds.
GPO Stock #017-024-01388-9 $10 NTIS PB #90-148941/AS $39

94 PHARMACOLOGY AND TOXICOLOGY OF AMPHETAMINE AND RELATED DESIGNER DRUGS. Khursheed Asghar, Ph.D., and Errol De Souza, Ph.D., eds.
GPO Stock #017-024-01386-2 $11 NTIS PB #90-148958/AS $39

IN PRESS:

95 PROBLEMS OF DRUG DEPENDENCE, 1989. PROCEEDINGS OF THE 51ST ANNUAL SCIENTIFIC MEETING. THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.

96 DRUGS OF ABUSE: CHEMISTRY, PHARMACOLOGY, IMMUNOLOGY, AND AIDS. Phuong Thi Kim Pham, Ph.D., and Kenner Rice, Ph.D., eds.

97 NEUROBIOLOGY OF DRUG ABUSE: LEARNING AND MEMORY. Lynda Erinoff, ed.

98 THE COLLECTION AND INTERPRETATION OF DATA FROM HIDDEN POPULATIONS. Elizabeth Y. Lambert, M.S., ed.

99 RESEARCH FINDINGS ON SMOKING OF ABUSED SUBSTANCES. C. Nora Chiang, Ph.D., and Richard L. Hawks, Ph.D., eds.

100 DRUGS IN THE WORKPLACE: RESEARCH AND EVALUATION DATA. VOL. II. Steven W. Gost, Ph.D., and J. Michael Walsh, Ph.D., eds.

101 RESIDUAL EFFECTS OF ABUSED DRUGS ON BEHAVIOR. John W. Spencer, Ph.D., and John J. Boren, Ph.D., eds.

102 ANABOLIC STEROID ABUSE. Geraline C. Lin, Ph.D., Lynda Erinoff, Ph.D., eds.

103 DRUGS AND VIOLENCE: CAUSES, CORRELATES, AND CONSEQUENCES. Mario De La Rosa, Ph.D.; Elizabeth Y. Lambert, M.S.; and Bernard Gropper, Ph.D., eds.

☆U.S. GOVERNMENT PRINTING OFFICE: 1991-286-729

DHHS Publication No. (ADM) 91-1722
Alcohol, Drug Abuse, and Mental Health Administration
Printed 1990